



EVALUATION

USAID/Dominican Republic Batey Community Development Project

September 2013

This publication was produced for review by the United States Agency for International Development. It was prepared by Hugo J. Pirela, Ph.D. & Emmanuel Silvestre, Ph.D., for USAID Dominican Republic under Contract No. AID-517-O-13-00002.

TABLE OF CONTENTS

	PAGE
EXECUTIVE SUMMARY	a
I. Introduction	1
A. Evaluation purpose & questions	1
B. Project background	2
C. Methods & limitations	3
1. Validation variables	4
2. Dependent variables	4
3. Statistical tools & tests	6
4. Procedure	6
D. Qualitative analysis	7
E. Prior data & measurements	7
II. Findings 1: Project achievements and observable transformations	8
A. Project outputs	8
1. Health	9
2. Education	11
3. Sanitation	12
4. Community mobilization	13
B. Project outcomes	15
1. Results Vs. targets, observable effects	15
a. Intermediate outcomes	16
(i) Health	16
(ii) Education	17
(iii) Sanitation	18
b. Terminal outcomes	18
III. Findings 2: Attributable results.....	23
A. Quantitative perspective 1: SPM-Hato Mayor bateys, phases I & II	24
1. Project outputs	24
a. Health	24
(i) Attention & treatment	24
(ii) Wellness promotion	26
b. Education	27
(i) Refurbishing of school infrastructure	27
(ii) Enhancement of teacher competencies & tools	28
c. Sanitation	29
(i) Revamping of sanitary infrastructure	29
(ii) Upgrading of water systems	32
2. Project intermediate outcomes	32
a. Health	33
b. Education	34
c. Sanitation	36
3. Project terminal outcomes	38
a. Health	38
b. Education	42
c. Sanitation	42

4. Other project results	43
a. Housing	43
b. Emergency preparedness	44
c. Income generation	45
B. Quantitative perspective 2: Other bateys and regions of phase II	46
1. Results in the region of Verón-Bávaro	46
2. Results in the region of Bahoruco	47
C. Overall qualitative perspective	48
1. Integrated approach to community development	48
2. Community organization & empowerment	49
3. Grassroots participation & self-management	49
4. Cross-sectoral enabling strategy	50
IV. Analysis & conclusions	51
A. Effect of the Project on the batey residents' quality of life	51
1. What was expected of the Project	52
1. Objectives attained	52
2. Changes confirmed	53
3. Results attributed	53
(i) Verifiable deliveries	54
(ii) Verifiable early effects	54
(iii) Verifiable advanced effects	55
2. What was not expected from the Project & the role of external factors	55
B. Effect on participation, organization & mobilization of batey communities	58
C. Effect in creating partnerships with key external actors	59
V. Recommendations (for possible future efforts)	59
ANNEX I: Original project theory	
ANNEX II: Note on the Evaluability of the Project	
ANNEX III: Research methodology	
ANNEX IV: Statistics (on physical CD)	
ANNEX V: Research instruments (on a physical CD)	

EXECUTIVE SUMMARY

The present study was commissioned by the USAID Mission in the Dominican Republic to assess the development effectiveness of the *Batey Community Development Project* (BCDP) funded by the Agency and implemented in the D.R. by *Save the Children* US; *Save the Children* Dominican Republic; and *Mujeres en Desarrollo* (MUDE) from November 2008 through June 2013.

The approved approach to the present evaluation was first to assess whether the project achieved or not its own stipulated objectives (*achievement* criterion). Other unintended or indirect project consequences, as well as the role of any external factor at play in shaping the actual end results were also to be studied, but only second to judging effectiveness against the Project's original intent. This brought to the fore the issue of *evaluability at entry*, i.e.: the extent to which performance could be judged against a pre-ordained set of measurable objectives, metrics and measurement procedures established at inception. The review of this issue yielded mixed results, as the requirements for evaluability at entry -in the internationally accepted definitions adopted herein- appeared only partially fulfilled in the Project's documentation. Designers clearly made an effort at the inception to define expected results. Also, at various points before the start, metrics were identified to gage such results. Yet, the structure and narratives in the project's original "Results Framework" did not quite conform to normalized notions accepted in the trade, including the fact that the identification of metrics showed a relative emphasis on the number of indicators, rather than on their quality and focus. Indicators were somewhat biased toward Project *outputs*, with fewer metrics on project development *outcomes* and none on project development *impacts*, and frequently did not include a baseline measurement against which to compare and analyze end line results. Also, strictly speaking, targets were not set "ex ante", but often in a "moving" pattern, separately defined by year. In view of these limitations, the researchers made use of the best available data and retrofitted the Project's development hypothesis (*project theory* or expected *Results Chain*) consisting of: **deliverables** (outputs) $\Rightarrow$ **changes of living conditions** in bateys (intermediate outcomes) $\Rightarrow$ **changes in resident's behavioral patterns** (final or terminal outcomes).

Other two central criteria used to assess the development effectiveness of the BCDP were: a) if there were observed changes in the batey's reality that were associates or concomitant in time with the Project's execution (*association* criterion); and b) whether the changes produced were statistically attributable to the Project (*attribution* criterion). Researchers were able to test the reconstructed hypothesis, applying these two criteria to few indicators in twelve intervened bateys of the *San Pedro de Macorís (SPM)-Hato Mayor (HM)* zone, based on empirical evidence gather from random samples of batey households, both by previous studies hired by Project's executors and by the present study's own field surveys. However: (i) not all indicators were measured in all bateys by previous studies, for comparison; (ii) not all indicators were provided baseline measurements from the start; and (iii) data on control groups were not always provided by previous studies. These facts severely restricted the full application of the two additional evaluation criteria mentioned above. Research in the SPM-HM zone was also affected by one or more of these field data restrictions; with the overall result that, the researchers were able to actually make meaningful quantitative comparisons only on a very limited number of indicators. between base-line and end-line figures, and between changes in Control and Treatment groups Also, whereas the random assignment of the control group for comparison is an internal consistency requirement of the study design stipulated by the Agency, the present researchers

did not participate in the randomization process used to select the groups at the start of the Project, and had to take the defined Control Group as a given for their field study

Finally, it must be also emphasized that the whole set of evaluation metrics set out at the start of the operation focused entirely on the all important question of: *to what extent did the BCDP Project affect the quality of life of the batey residents*. Other research questions stipulated by the Agency were left out of this quantitative evaluation framework; namely: a) *what effects did the project have on participation, organization and mobilization of the batey communities*; and b) *to what extent was the Project successful in creating partnerships with key external actors*. The quantitative testing of the above mentioned development hypothesis, which focuses on the first question and takes up the preponderant portion of this study, is later complemented with analysis of qualitative information, which was gathered through surveys of, and in-depth interviews with, Project executors, community leaders and resident field workers, mainly addressing the other two research questions

Quantitative findings & conclusions. The data reveals that in the Bateys of phase I –the most representative, because it included the majority of interventions- the Project reached or exceeded roughly two thirds of all its annual operational targets. Specifically 36% of annual targets were met on the mark and another 28% of targets were surpassed. This assessment gets somewhat qualified when performance is judged against final year (project-end) targets -arguably, a better criterion for overall operational success. The data reveals that 38% of final-year targets were not met in phase I of the Project. Also a wide difference is revealed when performance is assessed in terms of the submission of *development outputs* (project *deliverables*) as opposed to the assessment in terms of *development outcomes*, i.e. changes in living conditions and behavioral habits of the Project's target population, which were expected as a result of the intervention.

Based on the data cited above, the Project performance in terms of output delivery can be deemed strong. 77% of annual output targets in phase I were met on the mark or exceeded by the Project. Also, when the final year output targets –as mentioned: a stricter criterion of success- is considered for the assessment the Project registers an even stronger performance: with fully 83% of these targets met or surpassed in phase I. The less representative Phase II registers a similar delivery record, with 70% of annual targets and 80% of project-end targets attained or surpassed. Finally the generalized positive perceptions of community leaders and field participants in the Project, as gathered in the study's qualitative interviews, confirm the positive opinions of Project executors also interviewed, and attest to the quality of the operational products delivered for the benefit of the intervened Batey communities.

Researchers conclude that the performance above summarized constitutes a *prima facie* appropriate operational basis to expect the induction of “higher tranche” development results, imputable to the Project's actions. However, when measurement of development outcomes is taken as criterion for judging performance, the assessment is no so impressive. Less than half of the annual *outcome* targets set out (48%) was actually met or surpassed in the bateys of phase I of the Project. Outcome inducement appears to have been even weaker when end-project targets are taken into consideration: less than one third of all end-project *outcome* targets (29%) appear to have been achieved or surpassed by the Project in phase I. A much better performance on this respect is reported for Phase II; with only 30% of annual outcome targets, and 21% of project-end outcome targets not being attained. These results, however, imply only that quantitative targets were not attained, not that the Project might have

not made a significant difference in the living conditions and behaviors of the target population. Examples of these are the targets that were set about the *reporting by care givers of key children health and nutrition practices*; or the *identification by adults of STI and HIV/AIDS prevention practices*; or the *hand-washing habits of kids before eating at school*. In all of these areas visible strides were made, but the Project did not quite reach its own stipulated final year targets.

Concerning differences that can be time-associated to the Project in comparison with the ex-ante situation, only a fraction of the development outcome indicators defined included a baseline measurement. However, in this limited number of indicators the field data provided clear evidence that such changes did take place, and were concomitant with the execution of the Project. All documented changes were positive, with the exception of one, and spanned the gamut from health and wellness practices in adults and youngsters, to elementary school enrollment and graduation rates. The exception to the rule of advancement is the indicator of children enrolled in pre-school, which actually declined significantly in the SPM–Hato Mayor bateys. There is no lack of examples of dramatic positive changes as well, such as the one observed on the incidence of diarrhea in children under 5, in the bateys of the SPM–HM region, which dropped precipitously and concurrently with project execution. Also large increases in the value of indicators, which occurred concurrently with the intervention as well, have been documented in relation to children health & nutrition practices; women reproductive health practices; and STI, HIV/AIDS prevention practices.

In terms of whether or not the development results observed in the wake of the BCDP can be statistically attributable to the workings of the Project itself -which is the more stringent success criterion applied in the present evaluation- the evidence is also rather conclusive, although restricted to the limited number of indicators that could be so evaluated. For the reasons already expressed above, only a fraction of the identified indicators could be actually used in the full comparison between *treatment* and *control* groups. Nevertheless, the ones that could be compared yielded generally high values in the tests for statistical significance of the gaps observed and, thus, provided fairly conclusive evidence that the delivery of the Project's development outputs did induce some changes in the reality of the intervened batey communities. This evidence includes indicators that, although strictly referred to the *delivery of outputs*, could only reach noticeable values if a degree of cooperation, or change in attitudes, were present on the part of the beneficiaries themselves. This is the case, for instance, of the observed outputs about: *women receiving pap smear tests*; or *enrolled primary school students attending classes*. A variety of cultural or economic constraints, such as taboos about women's body or the frequent presence of child labor, traditionally inhibit or hamper these particular behaviors, in poor, uneducated communities. Therefore, the outputs corresponding to these indicators could not really be delivered adequately if those constraints did not somehow had already begun to crumble, and unless a modicum of disposition on the part of the population was present for the task at hand. Is the opinion of the present researchers that this type of budding behavioral changes may signal a trend in the direction of a more permanent and generalized future change in habits.

In fact, while the BCDP evaluation framework did not permit proper measurements of project impacts, (i.e. contributions of the Project to long term, structural changes in the bateys' reality) the study has documented some advanced development outcomes already visible by the end of operations. These had the form of incipient changes in behavioral patterns in the target population which certainly can be seen, especially if they are sustained, as precursor marks for longer term, more permanent

transformations, maybe coming in the future. These observed trends are time-associated and aligned with the development outputs delivered and the proximal development outcomes that appear induced by the BCDP, and they do not seem to be explainable by pure chance.

Researchers also found and discussed a few unintended results, denoting sometimes positive (desired) trends or consequences for the population's quality of life and sometimes negative (undesired) ones. Such results appear not to be random outcomes, yet at the same time may not be directly attributable to the Project but to other unexpected disturbing factors, perhaps at play. No hard evidence of the presence of such factors could be procured in any of the cases, but the study offers some hypothetical explanations for the intriguing results, associated with risk factors identified in the retro-fitted BCDP *LogFrame's* assumption column. The most conspicuous result, on the undesired side, is the precipitous decline registered in *DPT3 vaccination of infants* on the Control group, while the same vaccination in the Treatment Groups increased to the point of surpassing all annual targets in the same period. In this case, among the plausible external explanatory factors are: (i) a possible decrease in the scope of alternative government vaccination programs in areas not targeted by the Project; and/or (ii) the BCDP intense competition for a fixed and limited supply of officially controlled vaccines, in detriment of control bateys. On the desired side, a virtually identical trend is observed in the Control Group and in the Treatment Group -albeit with different base-line and end-line values- concerning the *PAP smear tests received by women* (upward) and concerning the *incidence of Diarrhea in small children* (downward). The hypothesis offered here by the researchers is that there might be positive "contagion effects" on behaviors between the compared groups, by virtue of role modeling or word of mouth; contagion which is possible because the groups selected were naturally porous and not completely impermeable to each other. This is, arguably, a non-random measurement error in the assignment of groups at the start of the Project, but one that is virtually impossible to always avoid when researching open human systems.

Finally it must be stressed that, normally, transformations associated with development outcomes require time to take hold; especially those already bordering the frontier with preliminary impacts, such as initial changes in behavioral patterns. So, in what concerns attaining targets, BCDP executors and promoters appear to have been self-defeating in, sometimes, setting unrealistically high expectations about such attitudinal and behavioral transformations only after a few years of induction and education; and with, perhaps, not enough regard for external factors possibly affecting results, such as the resistance that changes induced "from outside" always tend to generate in the targeted communities, given the prevailing culture and socio-economic reality.

Qualitative findings and conclusions. The study included in-depth interviews with project managers and executors on more qualitative aspects of the Project; as well as with community leaders, members and field workers on intervened communities of the SPM-HM and Verón-Bávaro regions. Such interviews were designed to validate issues of development approach, social methodology and quality of results; especially in what concerns the practical implementation of the Project's *cross-sector enabling strategy* and *integrated approach to community development* in the field.

A first important corroboration gleaned from systematic exchanges in the field with community leaders, ordinary residents and field worker is that, in all interventions, the Project did follow an approach to community development based on: (i) eliciting collective self-management and social mobilization; (ii) promoting community consensus on needs and necessary actions; and (iii) the practical study and application of ways to induce behavioral changes in individuals and groups. This field approach appears

to have been instrumental in strengthening grassroots organizations, social cohesion, awareness of development issues -especially among batey women- and community initiative in solving identified needs. The approach has emphasized grassroots participation at all points of the process, as well as a combination of: (i) specialized training on specific subjects -such as disease prevention, initial and vocational education, water systems maintenance, etc.; and (ii) the actual delivery of goods, services and tools for the benefit of residents -such as new or refurbished infrastructure for health; education; sanitation; emergency preparedness, etc. A salient point also stressed by the Project is that these delivered goods, services and tools were to become self-managed by the community for future sustainability. Such combined stress on training about development habits & values and on delivery of practical, material solutions also appears to have been successful in conveying the importance of social organization and mobilization to guarantee maintenance and sustain results; as well as for the general credibility and effectiveness of the Project.

With exceptions, researchers detected in most intervened bateys the presence of functional and active community organizations which were launched mainly due to the Project induction and promotion activities. These organizations included: local water boards; health committees; education committees; emergency committees; and women organizations; as well as organizations for the maintenance of the Community Center; youth clubs and energy committees. Members tended to know with precision what the organizations do and strive for in each case, and to speak fluently about their frequent meetings and activities. Residents and grassroots leaders also confirm the influence exerted by the Project in motivating and/or reviving social organizations of prior existence in the intervened communities. Often they bore witness to the fact that organizations which existed before the start of the operation, but were lethargic or not functional at all -such as residents associations; agricultural associations; or parents-teachers associations- were revitalized once the Project started and the message of integration and self-management began to be conveyed in the bateys. The Project's influence in empowering groups and individuals at the grassroots level is particularly visible in the workings of the committees for disaster mitigation and response, the local water boards and the women's groups, currently operating in the bateys under study.

Field interviews also reveal that the intervened communities remain fairly involved in developing viable solutions to their self-identified problems and needs. Collective action as the one displayed for community works during the Project in construction of the community centers, and house repair and refurbishing, is still frequent in areas such as cleaning and disposal of waste, as well as water system repair and maintenance. Residents and leaders report that this type of activities are more frequent now than in years before the Project, when only political mobilization in times of elections were mostly the norm, and repairs or re-constructions quickly deteriorated for lack of community attention and maintenance. An area of frequent mobilization often mentioned in the in-depth interviews is the all important maintenance function of the local water boards and the contribution of community labor to the revamping of water systems. Yet community actions in the pursuit of other goals are not infrequent, including collective petitions to authorities. A significant plurality of interviewed residents rated as *high* or *very high* the level of perceived participation of people in the affairs of the community, although some also rate it as *low*. Also, a majority of individuals interviewed rate the success of the community mobilization as *high*, though some in the minority rate such success as *low*. Even in some cases of communities with an appreciable level of mobilization, leaders report that the Project prepare them better for collective action, both in what concerned promoting the welfare of individuals and families

(such as through training on health preventive practices) and in promoting the welfare of the community as a whole. A particularly appreciated qualitative outcome of the BCDP in the intervened community is the universal agreement among interviewed community residents and leaders about the positive influence of the Project in lessening internal community violence and conflict. Some went as far as expressing that the BDCP all but eliminated the water conflict in their communities.

Beyond the palpable benefits recognized by the interviewed batey residents from the several interventions of the Project –particularly in the area of Health, where the role of the children *vaccination cards* and the work of the community health promoters is highly valued by people for their role in the so-called *community case management* and in promoting PAP tests and family planning, as well as the prevention of VIH, TBR, Dengue and Cholera,– benefits are also recognized and appreciated from the enabling strategy implemented by the Project in the intervened communities. Particularly noteworthy is the appreciation of residents for the emergency plans enacted by the Project, together with the provision of safe infrastructure and equipment in preparing for emergencies and responding to disasters. People spoke knowingly during interviews about the nature of such emergency plans and of specific cases in which they have been practically launched during the last year, or even within the last month at the time of the interview. Some mentioned, however, that the refuges are not appropriate or that they are already damaged.

An area raking lower in the general public appreciation of Project benefits is the area of *income generation* activities. Such activities are recognized as present and beneficial to people, both in what concerns family gardening plots to complement household nutrition and incomes, as well as in relation to small-loans granted for micro-businesses in the intervened bateys. Yet, many declare no knowing or being benefitted, especially from the latter activities. Finally, whereas key coordination with public institutions were clearly made to organize the delivery of certain project outputs, such as the launching of MISPAS health reference systems in the intervened communities, and the training of batey school teachers, based on contents of text approved by authorities of education, and notably with the private sector through the there is no clear evidence for a meaningful development of partnerships with public institutions involved in the bateys, as a longer term strategy to promote sustainability of results.

I. Introduction

1.1 The present study was commissioned by the USAID mission in the Dominican Republic to assess the development effectiveness of the *Batey Community Development project* (BCDP) funded by the Agency and implemented in the Dominican Republic from November 2008 through June 2013. The Project sought to induce sustainable improvements in the living conditions of the “*Bateys*”: former sugar cane work camps which are home to poor Haitian migrant workers and Dominicans. The Project aimed to focus on basic health, education services, income generating activities and linkages to other programs that can also contribute to provide livelihood improvements to said communities.

A. Evaluation purpose & questions

1.2 The study aims to measure the Project’s development effectiveness, analyze its implications and sustainability, identify lessons and provide feed-back to stakeholders on the Project’s operational and strategic guidelines. Specifically, the study seeks to answer questions on: (i) to what extent has the BCDP Project affected the quality of life of the batey residents; (ii) what repercussions has the project had in participation, organization and mobilization of the batey communities; and (ii) to what extent has the Project been successful in creating partnerships with key external actors, such as the Ministries of Health and Education. The evaluation premises and indicators approved before launching the BCDP focused on the first question; allowing the study to addresses issues with a detailed quantitative approach and a testable development hypothesis -within the limits of such premises and indicators. Questions ii and iii are addressed on a more qualitative section that complements the study findings.

1.3 Given the said important limitations and difficulties found with the evaluation premises and metrics of the present Project -as discussed in detail in the *Note on the Evaluability of the Project*- the basic development hypothesis that the study endeavors to prove has been retrofitted by the researchers, based on available Project information. The idea was to strike the best possible alignment between actual metrics at hand and the requirements of the methodological approach adopted.¹ The reconstructed development hypothesis is summarized in the

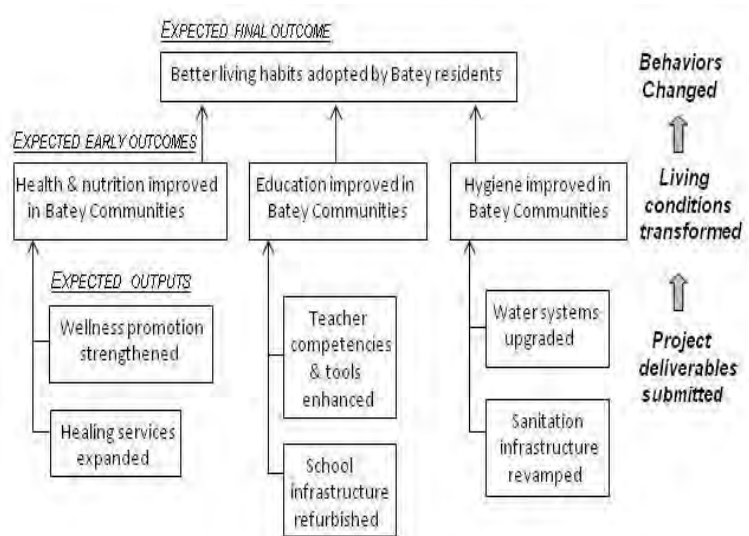


Figure 1.1: BCDP Development hypothesis (Retrofitted expected results chain)

¹ See the original “Results Framework” in Annex I, and the *Note on the Evaluability of the Project*, Annex II pp 5 - 13. Initially a *Foreword*, this note was moved to an annex at the request of USAID reviewers. Yet, we strongly suggest its early reading to clarify the technical limitations found and add perspective to the findings discussed.

detailed *results chain* presented in figure 1.1, which should be understood as the expected sequence of development objectives pursued by the BCDP (project theory).

- 1.4 This development logic, together with the full set of metrics identified for its verification, as well as some ex-post established risk factors, is included in the reconstructed Project *LogFrame* matrix on page xii of the *Note on the Evaluability*. In sum, the study answers the question about BCDP-induced changes in the bateys quality of life by measuring the extent to which the said hypothesis has been verified, in what concerns: (i) observable differences between: targets set and results achieved (*achievement criterion*); (ii) observable differences between base-line and end-line bateys' situation, that can be associated with the project (*association criterion*); and (iii) observable differences between the situation of treatment and control groups as they can be attributable to the project (*association criterion*)².
- 1.5 A few metrics not included in the *LogFrame* matrix, relating to project interventions in areas such *housing, emergency preparedness, and income generations* are also presently reported on. Yet results could not be reasonably associated with the integrated, measurable development logic presented herein. Finally, for reasons amply discussed in the *Note on Evaluability*, the development hypothesis and *LogFrame* matrix herein reconstructed represent a framework appropriate in theory for the evaluation of the BCDP; yet, the extent to which they could be thoroughly investigated was bound by the quality of field data and measurements provided to researchers. Disturbing factors stemming precisely from the way data was collected and measurements were done, prior to the present study, restrict the applicability of the theoretical evaluation framework and limit the scope of the present evaluation.

B. Project background

- 1.6 The BCDP was implemented by Save the Children International and two local NGOs (Save the Children Dominicana/FUDECO and Mujeres en Desarrollo Dominicana (MUDE), and was originally planned to benefit eight bateys of San Pedro de Macoris and Hato Mayor for a two-year period. In FY 2009, an additional \$4.7M was made available and allowed USAID to extend the implementation period from two to five years. It also allowed to strengthen BCDP activities (basic health care, nutrition, sanitation, and shelter for migrant workers and other residents), added additional bateys to the project in the surrounding region, and began expansion of some BCDP programming (specifically water and sanitation programming) to new geographic areas. One million of these additional resources were used in a partnership with USAID-Rotary International H2O Alliance (H2O Alliance) to bring water and sanitation access and improved hygiene and health conditions of batey residents.

² Note that the last two verifications can only be argued in terms of the pure differences in indicators between *Base-line* and *Project-end* measurements in the Treatment Group, as well as between Project-end measurements of the *Treatment Group* and those of the *Control Group*. Also, as discussed in par. 21 & 22 of the *Note on Evaluability of the Project*, arguments on whether the Project achieved or not its own objectives can be made only in terms of targets established annually, and not necessarily based on aspirations strictly specified *ex ante*, i.e. before the operation's start. Finally, the majority of the metrics defined before the start of the Project (about ¾ of them) are *output* indicators, where obvious differences are guaranteed to be found between treatment and control groups, simply by the fact that no products were delivered in the control groups. For a detailed explanation on the preponderance of output indicators (and not of those that truly measure *development outcomes*) and on the fact that the latter, on top of being scarce, suffer from various gaps encountered in the data made available to researchers, see the *Note on the Evaluability of the Project*.

- 1.7 BCDP partners coordinated with other USAID programs, particularly in health (HIV/AIDS prevention and treatment/ maternal and child health, and health systems reform), and education. Furthermore, the batey project expanded collaboration with other NGOs and Dominican government institutions such as *Consejo Estatal del Azúcar* (CEA), *Ministerio de Education* (MINED), *Ministerio de Salud Pública y Asistencia Social* (MISPAS), and *Instituto Dominicano de Recursos Hidráulicos* (INDRHI) and private organizations.
- 1.8 Public events were held in order to disseminate the project's results and encourage replication of the project model by the GODR, the Dominican private sector and other international organizations. The implementing consortium remained focused on the geographic micro-region (San Pedro de Macoris and Hato Mayor) where the BCDP was implemented; however, it included other geographic areas such as the Veron/Bávaro region in the east and Barahona in the South, where water and sanitation activities in partnership with the H2O Alliance were carried out.

C. Methods & limitations

- 1.9 The analysis and findings in this report result from the application of mixed research tools, including the analysis of quantitative data on project results, both already available and newly produced by the present researchers, in the context of a *quasi-experimental* study design which, to the extent possible, includes pre-tests and post-tests for both *treatment* and *control* groups.
- 1.10 Based on such design, and on the field data already available, a sizeable sample of not-intervened bateys and non-beneficiary batey residents (control group) was surveyed. The purpose was to measure changes in those control communities in terms of established *indicators* and compared them to the changes of the same indicators in the intervened communities (treatment group) measured by studies previously hired by the implementing partner *Save the Children*, both at the beginning (base-line) and at the end of the operation (end-line). This design corresponds to the #4 experimental design from Campbell and Stanley (1963); but in this case we deem it as a quasi-experimental design due to the lack of a real controlled randomization at the formation of the groups; i.e.: what was actually assigned to the Treatment and Control groups were the communities (Bateys), not randomly selected households.³
- 1.11 During the execution of the Project, the executors commissioned a firm to perform the Baseline and Final studies for the bateys on the Treatment Group and the baseline for the bateys on the Control

³ As differences in the performance of indicators between control and treatment groups might be attributable to the Project, the purpose of this study design is to find the maximum possible correlation between the end results observed and the means deployed by the Project; and, thereby, substantiate the strongest possible argument for attribution. It must be noted, however, that the requirements for this research model could not always be completely fulfilled for the present study, because of a number of data and design restrictions, beyond the said lack of randomization of groups; particularly: (i) that in some *treatment* and/or *control* communities no *base-line* and/or *project-end* study was done; and (ii) that it was virtually impossible for the groups to be impermeable and isolated to the point that could not influence each other's behavior, independently of the Project. Thus, frequently the researchers could not really come up with hard evidences of differences attributable to the Project; or of changes that might have occurred in the control group due to factors not controlled for. This and other practical limitations confronted by the quantitative study and analysis, along with the full set of the study's technical premises are discussed in the *Note on the Evaluability of the Project* and are spelled out in further detail in the annex on *Research Methodology* of the present report.

Group. The present researchers performed the Final study for the Control Group. Also the task of the present researchers was to compare the information gathered on the baseline studies from the firm that did them, with the information from the Final field studies, in order to evaluate the performance of the project indicators.

- 1.12 A sample size of 400 households was selected in the Control Group, as statistically representative for the final field survey. To this sample, whose distribution among the control bateys is specified in table 1.1, appropriate questionnaires reproducing the questions originally used in the previous studies were applied, to diagnose the current situation of these not-intervened bateys concerning the relevant indicators.

1. Validation variables

- 1.13 To validate the design analysis the researchers first controlled, to the extent possible, for the homogeneity of the household sample; i.e. made sure that the control households to be surveyed were indeed similar to, and valid for comparison with, those of the treatment group. To do this, several *independent variables* of the households were checked, based on the ones included in the original questionnaires; such as: sex, educational level and occupation of the respondent; number of household members; characteristics of the house, as the type of lightning, floor, walls and roof; and the kind of fuel used in the kitchen.

- 1.14 Another important validation performed on the sample chosen was to make sure that the communities selected for the Control Groups were in fact not intervened by the Project. For this part of the validation, researchers included questions in the Final surveys about knowledge of the existence of the Project and about having benefited from its activities. It is noteworthy that this validation process led the present researchers to discover that the Batey *Los Blocks de Mena*, officially included as belonging to the control group, during the execution had in fact been intervened by the Project; circumstance that, therefore, invalidated it as such control group.

2. Dependent variables

- 1.15 The dependent variables measured in the present study are included in the following list of indicators:

- a. **Health Area:** Improved maternal child, reproductive health, STD-HIV/AIDS prevention, TB prevention & treatment and hygiene, among batey residents.

of people benefitting from USG supported health services

% of children under five with diarrhea in the last two weeks

% children under 12 months of age who have received DPT3

% of children under than 5 years of age who received vitamin A from USG supported programs

SPM				
	BATEY	Households	Proportion	Sample (65%)
1	Olivares	110	0.35	71
2	Alejandro Bass	100	0.32	65
3	Los Chicharrones	80	0.26	51
4	Los Platanitos	20	0.06	13
	TOTAL	310		200
VERON-BAVARO				
	BATEY	Households	Proportion	Sample (4%)
1	Kosovo	1343	0.55	55
2	Matamosquito	627	0.26	26
3	El Macao	476	0.19	19
		2,446		100
BAHORUCO				
	BATEY	Households	Proportion	Sample (33%)
1	Mena Abajo	185	0.61	61
2	Los Blocks de Mena	119	0.39	39
		304		100

Table 1.1 : Samples households surveyed as final Control Group

% of children under 10 years of age who are de-wormed by USG-funded programs

of Community Health Promoters trained in:

- Maternal/newborn health
- Child health and nutrition
- Reproductive health/family planning
- TB management *

of Community Health Promoters equipped

of bateys with functional MISPAS reference systems*

% of care-takers who report at least three key positive child health and nutrition practices

% infants under 6 months of age exclusively breastfed within the past 24 hours

% women between the ages of 15 and 49 receiving at least one PAP smear test in the past 12 months

% of women between the ages of 15 and 49 who can report at least two key reproductive health practices

% of heads of households who identify at least key two STI and HIV/AIDS prevention practices

of adults tested for TB

% school-children who wash their hands:

- After using the school latrine/bathroom
- Before eating the school breakfast/snack/lunch

b. **Education Area:** Increased access to and improved quality of primary school education and extracurricular educational services.

classrooms rehabilitated and equipped

of classrooms with improved didactic materials

of teachers from participating batey schools trained in:

- Literacy
- Math
- Computer use
- School Health and Nutrition
- Other pedagogical improvements/practices

of administrators and education officials trained

% of enrolled primary school students attending

% of students enrolled in 4th grade that pass 4th grade

% of students in single grade primary schools in 1-4 grade who read at grade level*

of children enrolled in pre-school

of children and youth (6-14 years old) enrolled in USG supported extra-curricular programs

of youth and adults participating in USG-supported vocational or informal education programs

c. **Infrastructure Area:** Improved water, sanitation, school, and housing infrastructure and services.

of families with access to improved sanitation services

- # of families benefitting from improved community water systems*
- # of families with improved access to clean drinking water
- # families with improved housing

Enabling Strategy: Participatory Community Mobilization

- # of bateys with access to an emergency-safe structure
- # of bateys with updated emergency plan
- # of batey schools with updated emergency plan*
- # of residents receiving support in income generation activities*

(*) New indicators added in 2011, for the extension of the project

3. Statistical tools & tests

1.16 With an “a priori” *Power Analysis* researchers determined the total minimal sample required to ensure adequate *Size of the Effect* and *Power* of analysis for the study; thus ensuring that conclusions are not due to chance or random errors (i.e. false positives or false negatives of correlation). Through the said “a priori” *Power Analysis* the researchers made sure that the security standards usually applied in statistics were in place for the present study, namely: the probability of accepting a false positive was set at .05 and the probability of avoiding a false negative was set at .95.

1.17 Because the dependent variables in the present study must be measured on a nominal scale (categories and frequencies) the appropriate non parametrical statistic to be used for establishing significance of results is the *Chi-square* (χ^2) test. For this kind of approach, the *power analysis* takes into account the degrees of freedom of the comparison, so that higher the degrees of freedom imply greater minimal sample size required. Our basic comparison (the one in the SPM-Hato Mayor region, where the two project “phases” took place in full) had to be modified to include the fact that only one baseline study was available for the Control Group in the two Phases. Table 1.2 shows the modified comparison including the total samples used in the different studies. The Chi-square for this comparison has two degrees of freedom and, according to the result of the *power analysis*, the total sample size needed for an appropriate statistical analysis of our comparisons is 172 cases. As we can see in the above table, the other cells in the SPM-Hato Mayor comparison totalize 1,500 cases. Therefore, we were really free to select, as we did, a random convenience sample of 200 for the Final Control Group survey in SPM, with a total sample size of 1,700 cases: way above the necessary. A similar approach was used to choose the sample size in the other regions, for the total 400 households sample.

SPM Region	BASELINE	FINAL
PHASE I	304	461
PHASE II	342	209
CONTROL	184	200

Table 1.2 : Sample households, SPM region

4. Procedure

1.18 The sample of households was randomly selected in the Control Group bateys, excluding the commercial housings. The interviews were directed to key informants (household heads) of 18 years of age or more. The information on reproductive health practices was directed to women from 15 to 49 years of age in each household. The survey personnel was trained in administering a questionnaire adapted from the survey designed and used in the previous studies of the project.

D. Qualitative analysis

- 1.19 The study also obtained and analyzed qualitative information from surveys to community leaders as well as in depth interviews with project Executors, designed to glean specialized points of view on project performance; though always avoiding conclusions exclusively based on anecdotes and mere hearsay

E. Prior data & measurements

- 1.20 As it was note before, that the Project included a “second phase” or initiative that was not foreseen at the inception, with new target regions and bateys added “on the move”.⁴ The *intervened* regions and *communities* as well as the designated *control communities* and the availability of the field studies and samples of households surveyed in each by the previous studies, are summarily identified in table 1.3. Since separate base-line studies were envisaged for the new target populations, the fact that new regions and communities were added unexpectedly did not need be of major concern, because changes in the population targets of the second “phase” could still be presented separately (i.e. not in an integrated, unified way with results of the original first “phase”). However, not all indicators were applied to treatment groups in all bateys of the 2nd phase; so the chosen metrics could not be uniformly studied for all target populations. This was specifically the case of Bateys in the newly included regions of *Verón-Bávaro* and *Bahoruco*, where very few metrics were measured in the base-line study.

Phase I					Phase II					Phases I & II		
Treatment Group					Treatment Group					Control Group		
Regions/ bateys	Previous Studies				Regions/ bateys	Previous Studies				Regions/ bateys	Previous Studies	
	Pre	Sample	Post	Sample		Pre	Sample	Post	Sample		Pre	Sample
SPM-Hato Mayor					SPM-Hato Mayor					SPM-Hato Mayor		
Don Juan	✓	52	✓	94	- Euskalduna	✓	23	✓	77	- Olivares	✓	29
Cachena	✓	31	✓	39	- Margarita	✓	75	✓	47	- Alejandro Bass	✓	56
Prudencio	✓	16	✓	35	- Paraíso	✓	78	✓	47	- Chicharrones	✓	80
Consuelito	✓	50	✓	68	- Victoria	✓	66	✓	38	- Plataniitos	✓	19
Jalonga	✓	53	✓	82								
Dofia Lila	✓	16	✓	22	Verón-Bávaro					Verón-Bávaro		
Experimental	✓	45	✓	65	- Villa Plywood	✓	252	×	-	- Kosovo	✓	139
AB-4/AB6	✓	41	✓	56	- Hoyo de Friusa	✓	115	×	-	- Matamosquito	✓	64
										- El Macao	✓	58
					Bahoruco					Bahoruco		
					- Batey 3	×		×	-	- Los Bloks de Mena	×	-
					- Mena	×		×	-	- Mena	×	-

Table 1.3: Treatment & control groups in phases I, II. Sample: households studied; ✓: available; ×: not available.

- 1.21 Also, as indicated on Table 1.3, not in all bateys of the second phase a *base-line* and/or *final study* was done on the *treatment group* and/or on the *control groups*. In the treatment bateys of *Verón-Bavaro* no post-test was available at all, rendering impossible any project-end comparison with results in the control group obtained by the present researchers. For the case of *Bahoruco*, no *base-line* or *project-end* study was done in any batey of the *treatment group*, and no *base-line* study was done for the *control*

⁴ The designation “phase” is a bit misleading, because what actually happened in the project execution was that in 2011 a number of new bateys (4) and zones (2) were just added to the interventions already underway since 2009 in eight communities and one region; with both groups of intervention geographically overlapping and running parallel to each other since then, until they both finished at the same time in 2012.

group either. Finally, validation tests applied by researchers on the available data revealed that in the case of a *Bahoruco* bately which was deemed part of the *control group*, it actually could not be treated as such control group because some intervention had in fact occurred in that bately. All these perturbation factors, stemming from the way data was collected and metrics applied prior to the present study, resulted in the impossibility to produce a uniform evaluation for both phases of the Project, based on an integrated framework and a unified, quantitatively testable development hypothesis. This is especially true for the new regions of Phase II, where the perturbation facts render impossible to fully apply the framework. Yet, researchers were able to test the hypothesis in the region of *San Pedro de Macoris-Hato Mayor*, albeit on just a few indicators for development outcomes. This region includes the majority of treatment and control bateys; so general conclusions can be reached from the exam of results in this region alone. Accordingly, the preponderance of findings presented herein refers to that region in phase I and the rest will be presented in a separate section

II. Findings 1: Project achievements and observable transformations

A. Project outputs

- 2.1 Since, generally, the baseline for the operational deliverables is admittedly zero and no error can be presumed in attributing the submission of outputs to the project itself, we obviate the *association* and *attribution* criterion in evaluating project outputs. Thus, in what follows, we concentrate only on whether the project passed the *achievement criterion* of evaluation, i.e. if it reached or not the targets it set for itself, both annually and by the end of the operation (final year). Accordingly, this section only analyses the results presented in the executors' own operational reports -as summarized in the *Annual Performance Report FY 2012* (APR2012) from Save the Children.
- 2.2 Considering its annual targets taken as a whole, the project shows a strong performance in terms of

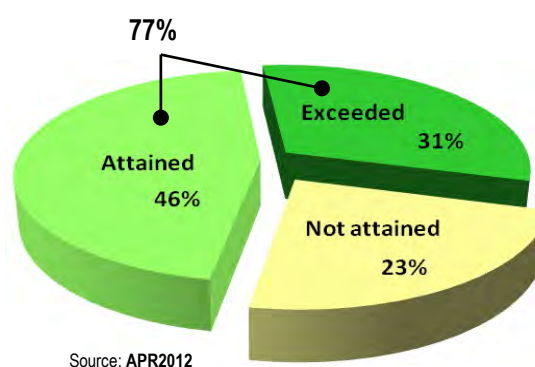


Figure II.1: Achievement of BCDP annual output targets

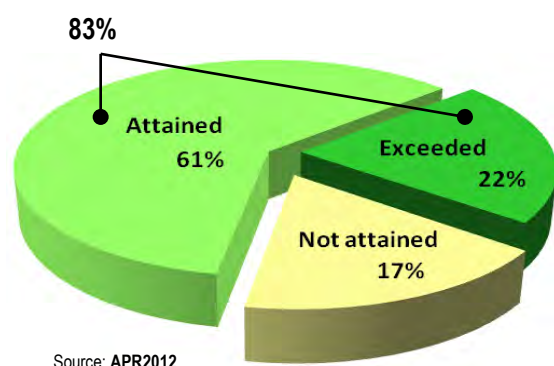


Figure II.2: Achievement of BCDP end output targets

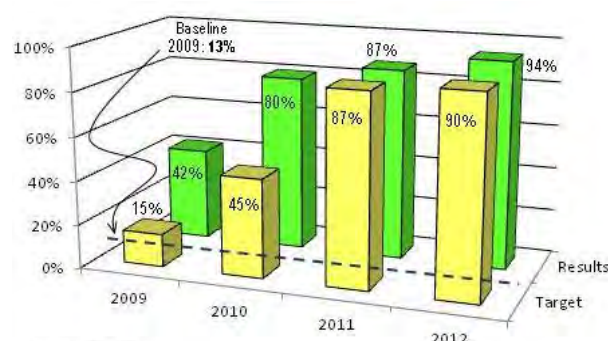
delivering its promised output. It reached a 77% achievement mark, if both attained and exceed targets are added together, as depicted in figure II.1. More importantly, when judged against final year (project-end) targets -arguably a better gage of the performance as a whole- the project shows even better marks. When this criterion is used, we discover that the Project attained or exceeded a full 83% of all its targets (See figure II.2) It is also interesting to note that this operational performance -i.e. performance in terms of deliverables submitted in time and form- is

also generally better than the project's total performance, i.e. the performance measured taking into consideration all targets, and not only those for the output tranche of the results chain.

2.3 Indeed, when the whole results chain is considered, data shows that the Project attained or achieved only 66% of all its annual targets, and only 63% of all its project-end targets. A detailed discussion of the delivery of outputs in the different components of the project is presented below.

1. Health

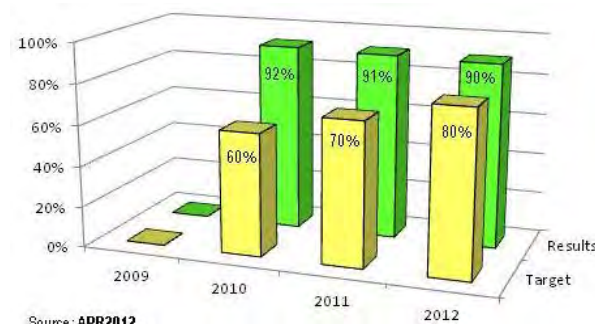
2.4 Figures II.3 to II.9 compare output target of the BDCP with annually reported delivery in the area of health on the 8 communities intervened in phase I of the Project. Beginning with vaccination and de-



Source: APR2012

Figure II.3: % of children under 12 months receiving DPT3

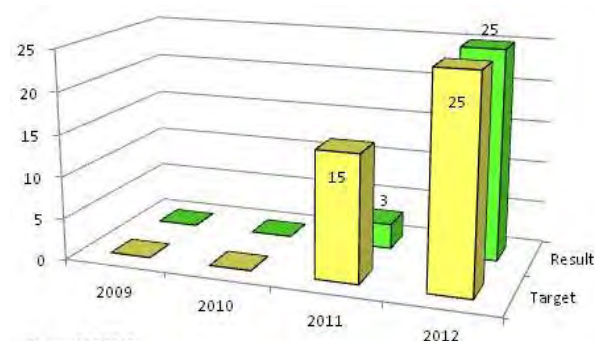
included a base-line value for 2009; i.e. the percentage of children of that age that received DPT3 then, presumably from other sources. As depicted in figure II.3, such proportion was significantly increased from 13% at the baseline to 94% by project-end in the intervened communities. The targets set for the de-worming of children under 10 years of age were also consistently exceeded throughout the execution of the project, with the sole exception of the first year (2009) where no de-worming has been reported, but no target was set either for such output (See figure II.4).



Source: APR2012

Figure II.4: # of children under 10 who are de-wormed

2.5 Results on the number of adults that were tested by the Project for TB were more mixed, as can be appreciated in figure II.5. First, no target was set for this indicator in the years 2009 and 2010 of the project execution. Then, actual results for 2011 fell way below the target for that year, and the annual target was just achieved for the final year 2012. Now, insofar as the annual targets set by project executors are understood to be “cumulative”, the important point herein is that in the final year (2012) the target set was actually achieved, despite the fact that performance was poor in the year 2011.



Source: APR2012

Figure II.5: # of adults tested for TB

2.6 The three indicators discussed so far can be considered measurements of actual health services provided by the project to inhabitants of the batey

communities intervened, as opposed to other activities more in the line of *wellness promotion* and *nutrition*, which were also deployed by the operation within the health component, and which will be discussed next in the present report. Now, in general, the targets concerning the *health services* proper, - both annual and project-end- were all exceeded or achieved throughout the execution of the project.

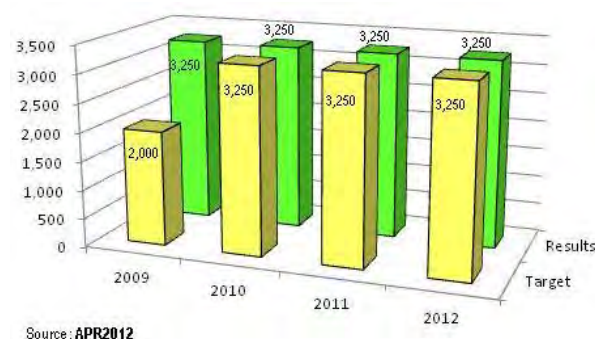


Figure II.6: # people benefitting from Health services

2.7 This can be seen in figure II.6, which depicts the number of batey residents that benefitted one way or another by the health services provided by the project in the intervened communities of the Project's phase I. The researchers found that the numbers reported by management on the indicator: *# of people*

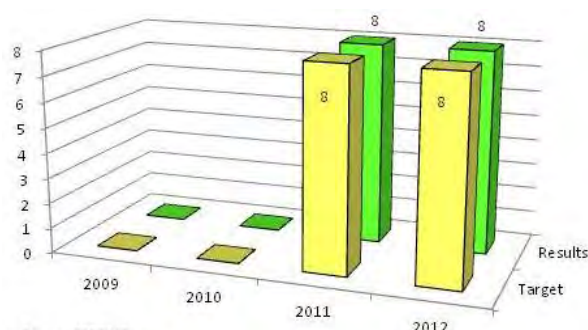


Figure II.7: # of bateys with MISPAS reference systems

benefitting from *USG supported health services* entail a degree of inconsistency, because while the stipulated definition of the indicator refer only to health services, the measurement actually made referred to people benefitting from all services provided by the Project. However, these figures would tend to substantiate the claim that, concerning *health services*, the output required to induce the expected health outcomes were generally put in place by the project. The same can be said of the wellness promotion/nutrition line of

activities where targets have been largely reached. Executors report the deployment of the official Health Ministry (MISPAS) reference systems, and the training/equipping of community health promoters. On the launching of the reference systems, Management reports to have achieved the target of eight such systems in the communities of phase I. (See figure II.7). The exceptions are years 2009 and 2010 for which no target or results values were reported. Although the implementation of the systems occurred relatively late, they may have induced some positive "contagion" in not targeted bateys, as we will later discuss. In equipping and training health promoters (on maternal/newborn health, child health & nutrition, reproductive health & planning, etc) Management reports to have prepared 17 individuals in the 8 communities intervened, which was the target value for both indicators on all years, with the exception of 2009, when the target was set at 15 and, therefore, exceeded. (See figures II.8, 9)

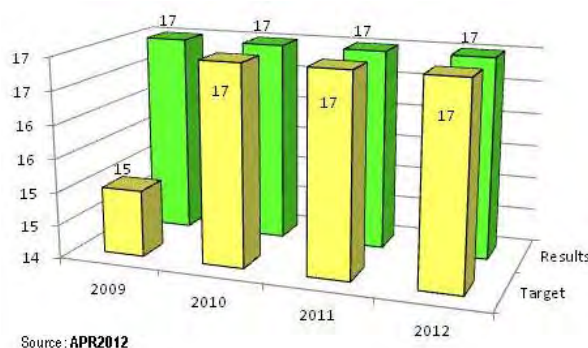


Figure II.8: # of community health promoters trained

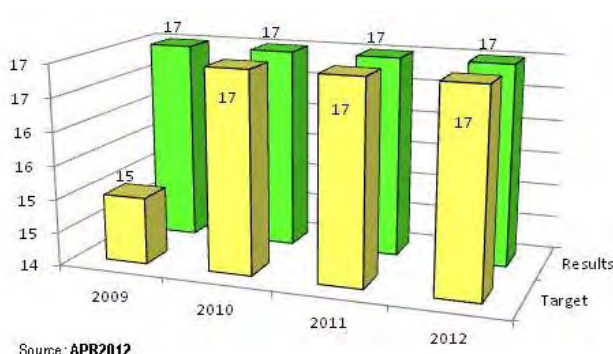


Figure II.9: # of community health promoters equipped

2. Education

2.8 In relation to educational outputs, figures II.10 and II.11 depict the results reported by Management in terms of classrooms rehabilitated and equipped, as well as those provided with improved didactic materials. The target of 22 classrooms for both indicators were reportedly reached every year, with the exception of year 2009; when the target was set at 5 classroom, and none was rehabilitated, equipped or provided with improved materials.

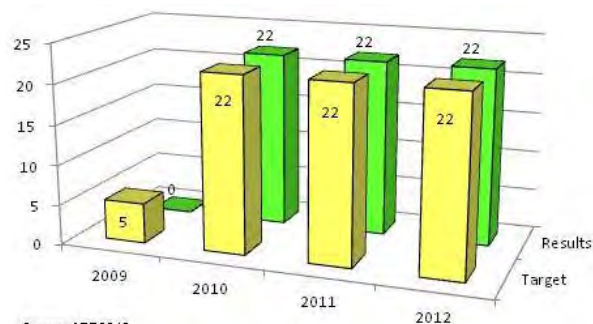


Figure II.10: # of classrooms w/improved didactic materials

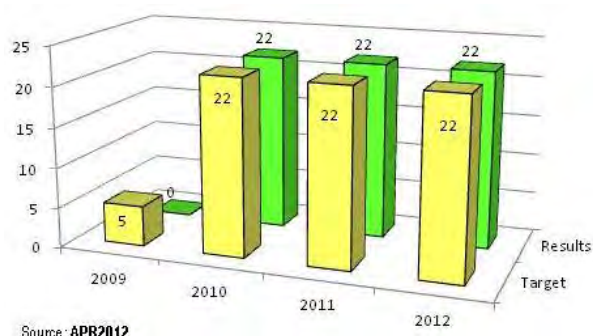


Figure II.11: # of classrooms rehabilitated and equipped

2.9 Results on the training of teachers in batey schools are mixed, as shown on figures II.19 – II.22 about relevant indicators on the subjects of: (i) Literacy; (ii) Math; (iii) Computer use; (iv) and School health and nutrition; indicators that appear to represent not cumulative but independent yearly values.

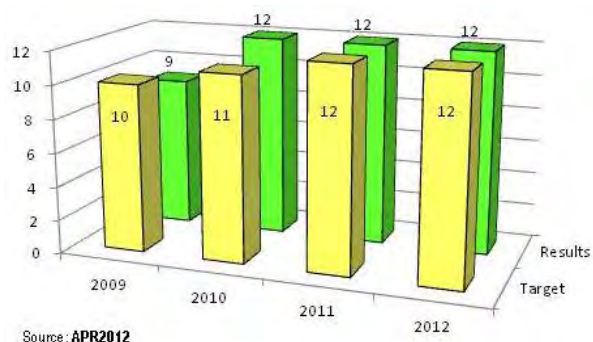


Figure II.12: # of batey school teachers trained in Literacy

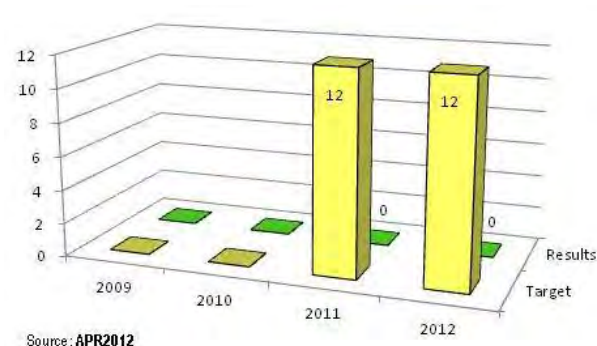


Figure II.13: # of batey schools teachers trained in Math

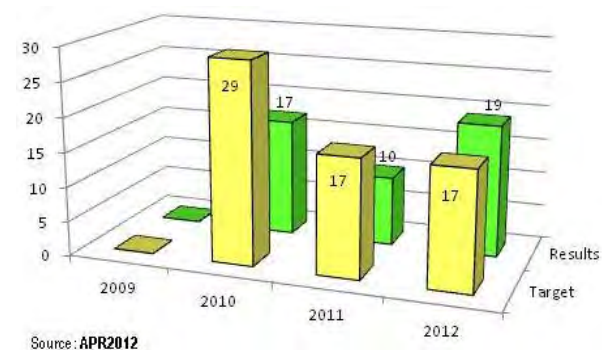


Figure II.14: # of batey schools teachers trained in computer

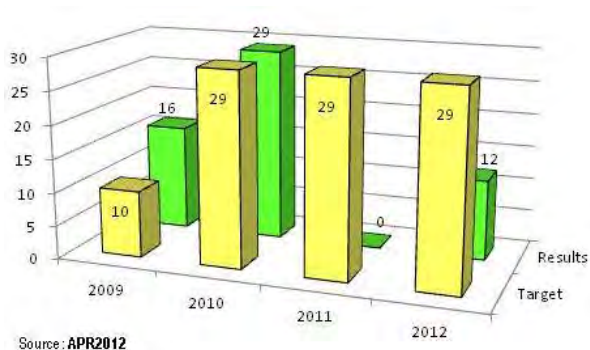


Figure II.15: # of batey schools teachers trained in school health & nutrition

2.10 The data shows that the Project was uneven in achieving its objectives for the training of batey schools teachers. After running the first two years with no target set out or actual results, the indicator of training on *Math* ended up yielding no results, even in years 2011 and 2012 where targets were stipulated. The reasons reported for this performance is that the time dedicated to training in literacy was initially cut short, and then had to be extended to complete the program, leaving no time for the Math training program. Management has pointed out that teacher strikes and other labor confrontations revolving around the adoption of new official textbooks by the national authorities caused the initial severe interruption in the literacy training; the same that eventually affected negatively the training on Math.

2.11 Results on other areas of training were also not up to par with expectations, as the case of training in *Computer Use*: area where the annual target was only achieved -and surpassed- on the final year of the Project, 2012. In the case of the training of batey school's teachers on *School Health and nutrition* targets were achieved or exceeded in the first two years, but results fell seriously below targets during the following years 2011 and 2012. Only the training of teachers in *Literacy* achieved its expected results, with the exception of year 2009, when the target was not reached.

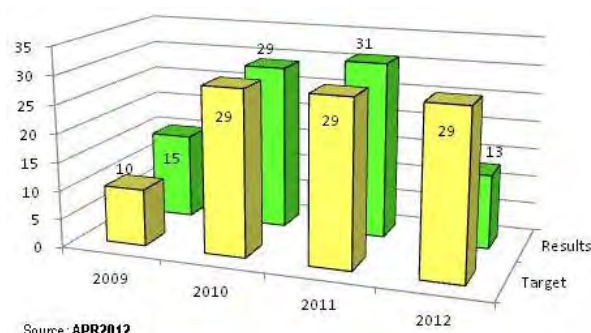


Figure II.16: # of batey schools teachers trained in other pedagogical improvements & practices

2.12 Another case of less than optimum project performance was that in teacher training on the subject of *Other Pedagogical Improvements and Practices*. The annual targets on this indicator were achieved or exceeded during the first three years, but fell way below target on the last year of the Project, as depicted in figure II.16. Finally on the area of training, the Project did achieve its stipulated annual targets in teaching school administration and education officials on subject matters that will help them better discharge their responsibilities.

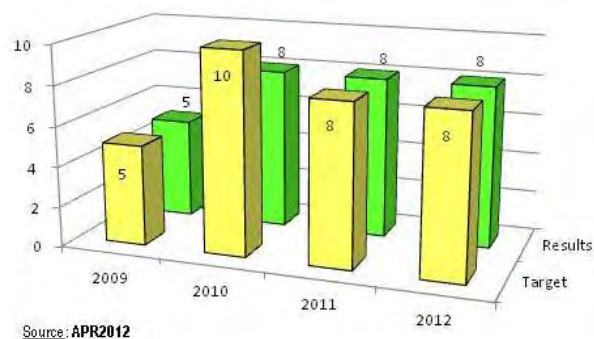


Figure II.17: # of administration & education officials trained

So, according to the performance just reviewed and contrary to what the health output indicators showed, a relatively sober performance of the Project is seen in what concerns the delivery of its promised educational outputs.

3. Sanitation

2.13 The indicators presented by Management as measures of development outputs delivered by the BCDP on the area of community hygiene and sanitation were: (i) # of families with access to improved sanitation services; and (ii) # of families benefitting from improved community water system

2.14 Figure II.18 illustrates results in terms of access of families to improve sanitation services in the intervened batey communities. In 2009 and 2010 the annual targets of 50 and 300 families respectively were not reached. For 2011 the annual target was reportedly lowered, and results exceeded it. Finally the last year of the operation the target of 300 families covered was achieved. Here again it is worthwhile to note that, provided that these annual targets were cumulative, the important fact is that the project-end target (2012) was indeed achieved.

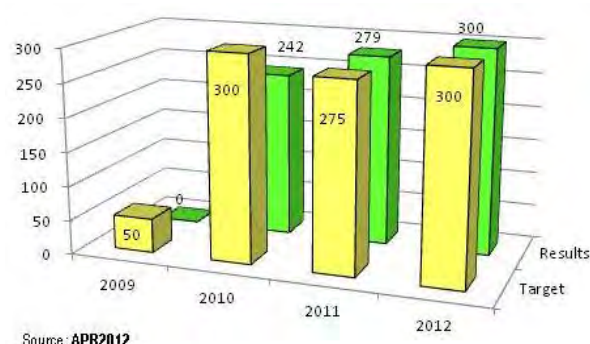


Figure II.18: # of families with access to improved sanitation services

2.15 Now, in terms of the number of families benefitting from improved waters systems, that number reached the final output level of 953 already in the second year of execution, although up to that point

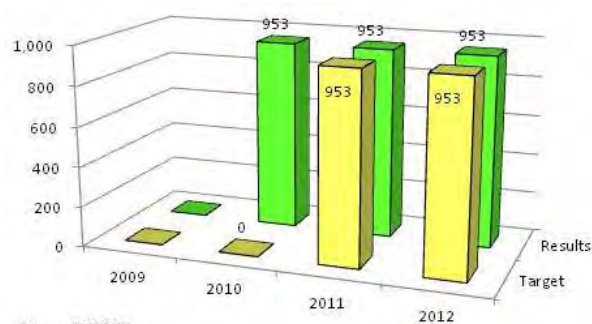


Figure II.19: # of families benefitting from improved community water systems

no annual target had been stipulated by project executors. See these results in figure II.19. The achievement of the targets set out for this particular component as whole arguably shows that the output requisites necessary were put in place by the Project, to induced the expected development outcomes up the results chain in the area of sanitation. Moreover, these outputs have arguably also important implications for outcomes in terms of the batey resident's general health and wellness, in a sort of expected "cross-component effect".

4. Community mobilization

2.16 Although community mobilization is understood -as conveyed in the original project documentation- as a cross-sectoral enabling strategy to promote all results, rather than a result area, the executors' report presents it as a result area itself, with performance indicators and annually planned targets and achievements. The indicators are: (i) # families with improved housing; (ii) # of bateys with emergency-safe structure; (iii) # of bateys with updated emergency plan; (iv) # of batey schools with updated emergency plan; and (v) # of residents receiving support in income generation activities.

2.17 In what concerns housing, after the first two years of falling behind schedule, in its last two years the Project, on the contrary, surpassed the respective annual targets and the intervened communities ended up with 600 families -50 more than the planned number- benefitting from improved housing. This included repairs as well as major refurbishing and construction, and constitutes a major socio-economic advancement for the poor families involved. See results in figure II.20.

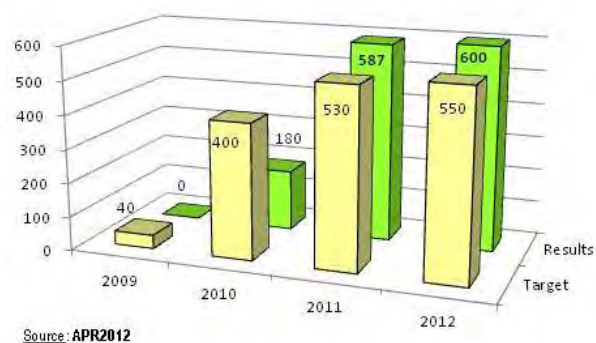


Figure II.20: # families with improved housing

2.18 The indicator on availability of emergency-safe structures was the only one in this series to include a baseline value; patently allowing a quantitative association of the Project with a change in the safety of

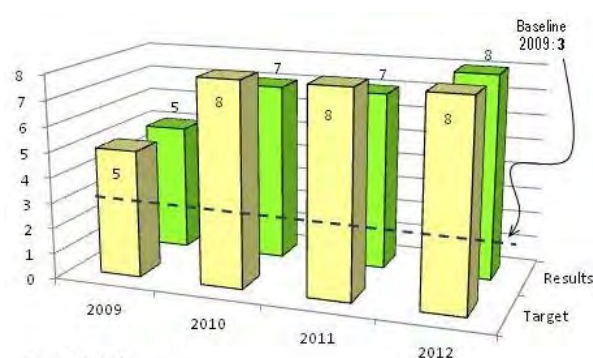


Figure II.21: # of bateys with emergency-safe structure

batey residents who benefitted from this effort. In this case, three of the intervened bateys had access to such a structure before the start of the Project. The first annual target -increasing this number to 5 batey communities- was achieved by the end of the first year of operations. Also, even though the objective of providing a proper emergency refuge to all 8 intervened communities was not reached in the two subsequent years, it was finally achieved by the end of the Project as illustrated on figure II.21. The Project also achieved visible concurrent results in

what concerns emergency preparedness of the intervened communities and schools. As can be appreciated in figure II.22, the Project goal of having all 8 intervened bateys in possession of a community-managed emergency plan was achieved from the very first year of operations. Also figure

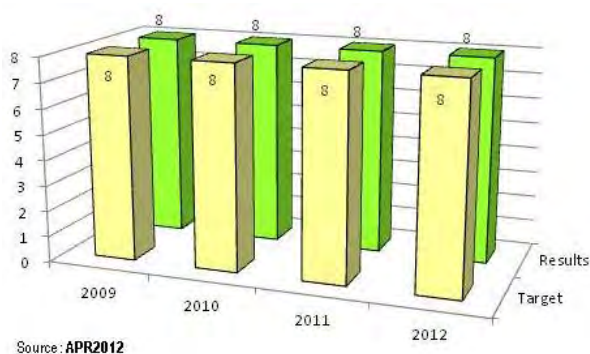


Figure II.22: # of bateys with updated emergency plan

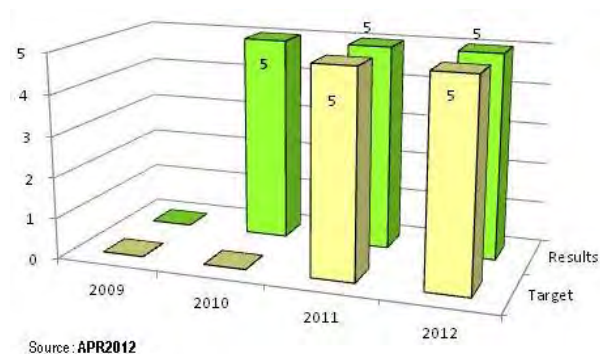


Figure II.23: # of batey schools with updated emergency plan

II.23 illustrates how the corresponding goal of having 5 batey schools with self-managed emergency plans was reached already in 2010, at a point on which no target had apparently been set out yet for this indicator.

2.19 Finally, Management reports results in the area of income generating activities, consisting mostly of home gardening to produce enough for family consumption -thus improving nutrition- and also for sale -thus increasing family income. Income generating activities also included the provision of credit for micro-businesses in the communities. Having not set expected goals for the first year, 2009, the Project reached its targets in the intervening years of 2010 and 2011; although, at the end slightly fell below the final target number of 225 residents participating in such activities, as can be seen in figure II.24.

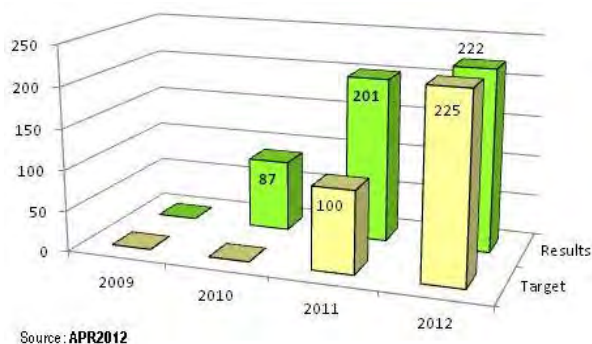


Figure II.24: #of residents receiving support in income generation activities

B. Project outcomes

2.20 As in the case of project outputs, the *achievement criterion* can be applied in evaluating the BDCP outcomes, both intermediate and terminal. And the *association* and the *attribution* criteria are also usefully applicable in the case of outcomes. Therefore, results of the Project are evaluated below on the basis of the three criteria. We start by examining whether or not the targets set by the Project were achieved (*achievement criterion*). Yet, in doing so, we will also apply the *association* criterion whenever possible; i.e. to those indicators for which a base-line was originally measured and, therefore, a provable change in reality can be time-associated with the workings of the Project. Then, on a separate chapter we will take up the issue of whether the observed changes could be attributed to the intervention (*attribution criterion*) on all indicators for which this investigation is presently feasible. The above analysis will be performed both for intermediate and terminal project outcomes in the several components of the BCDP.

1. Results Vs. targets, observable effects

2.21 As presented by Management in its own reports and considering the annual targets taken as a whole, the Project shows a much weaker performance in reaching the *outcome* objectives it set out to achieve, compared to the results shown to have been achieved in delivering project outputs. In fact, more than half of the annual outcomes targets were not attained by the project, as can be seen in figure II.25. By comparison, 48% of these targets were actually exceeded.

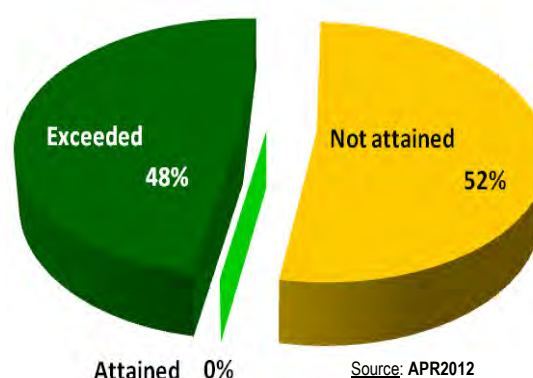


Figure II.25: Achievement of BCDP annual outcome targets

2.22 Even more telling, when final year (project-end) targets are taken into consideration to evaluate results - which is admittedly a better gage of the intervention's achievements as a whole- the project shows an even weaker performance. Actually, because outcomes are *effects* of transformation in the target reality that by definition tend to appear toward the end of interventions, the establishment of early annual targets is not generally appropriate for outcomes. For this reason, project-end targets would be commonly more adequate on the whole, to measure effects achieved. Now, when this latter criterion is used in evaluating the BCDP, the fact is revealed that the Project did not attain its outcome targets in more than 70% of the cases. (See figure II.26). These results are also below the average performance of the Project in terms of total annual and project-end targets; i.e.: the targets of both outputs and outcomes taken together for each one of the years of execution, including the closing year.

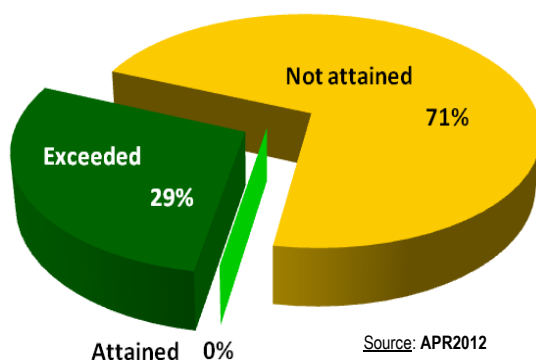


Figure II.26: Achievement of BCDP project-end outcome targets

2.23 A detailed discussion of the performance in terms of intermediate and terminal outcomes in the project components of Health, Education and Sanitation is presented below.

a. Intermediate outcomes

(i) Health

2.24 According to our working *LogFrame* matrix, the indicators available to measure early development outcomes of the BCDP in the area of health are: (i) *% of women ages 15-49 receiving one PAP smear test in the last 12 months*; and (ii) *# of children under 5 receiving vitamin A*. Admittedly these indicators are, strictly speaking, terminal output indicators; however, for the purpose of the present analysis the researchers accept that both fall near the frontier between *terminal outputs* and *early outcomes*, since some degree of attitudinal change and collaboration from the part of the beneficiaries is deemed necessary for these outputs to be delivered adequately; especially in the case of the women's PAP smear tests, around which a number of cultural restrictions relating to feminine intimacy taboos -especially on the part of husbands and other women's partners- were observed to operate as strong resistance forces against the project actually achieving its annual and project-end delivery targets.

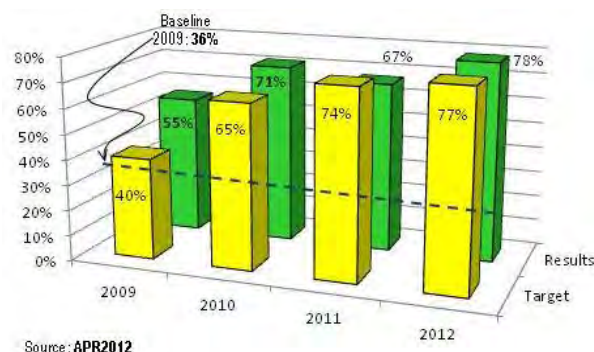


Figure II.27: % of women ages 15-49 receiving one PAP smear test in the last 12 months

2.25 In the case of this indicator an actual baseline value was measured at the start, and so a reasonable judgment can be passed on whether the Project may or not be associated with an observable change.

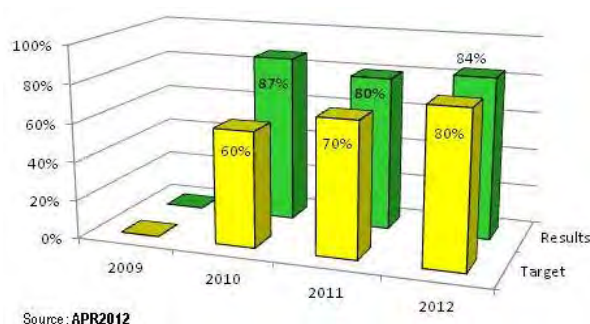


Figure II.28: # of children under 5 receiving vitamin A

Figure II.27 depicts how this indicator registered above target each year of the project, with the exception of year 2011. On that year, the result fell below the annual target of 74% and also below the previous year result of 71%. In fact, at a level of 67%, the 2011 result barely kept above the previous year target. Nevertheless, given the project-end achievement, the intervention can be associated with a total improvement of 42 percentage points in the indicator, above the base-line.⁵

2.26 Similar performance is reported in what concerns the provision of vitamin A to children under-5 years, as depicted in figure II.28. Except for year 2009, in which no target or result value was made available for the indicator, the Project reportedly exceeded all subsequent annual targets. In this case, unfortunately, no base-line value was made available and, therefore, no change with respect to the ex-ante situation can be discerned or associated to the Project.

⁵ Following our research premises and presentation order, whether this change -as other to be discussed- may actually be attributed or not to the project is a question that we will not address until the next chapter, when the appropriate statistical significance test are applied to the differences observed between results in the *Treatment* and *Control* groups. However, as in other cases from the numbers analyzed in the present section, at least a time-association can presently be argued between the execution of the Project and the changes observed.

(ii) Education

2.27 The indicators available to measure early development outcomes of the Project in the area of Education run the gamut from enrollment in pre-school, through attendance to primary school and enrollment in extra-curricular programs, up to participation in vocational education. Here again, although they fall very close the frontier with terminal project outputs, these proxy measures of improved education in the bateys arguably do reflect initial changes in conditions that may, or may not, point toward new behavioral trends in the population. Now, these budding trends appear no to have materialized in the case of *number of Children enrolled in pre-school*; an indicator that, thankfully, did include a base-line measurement at the start.

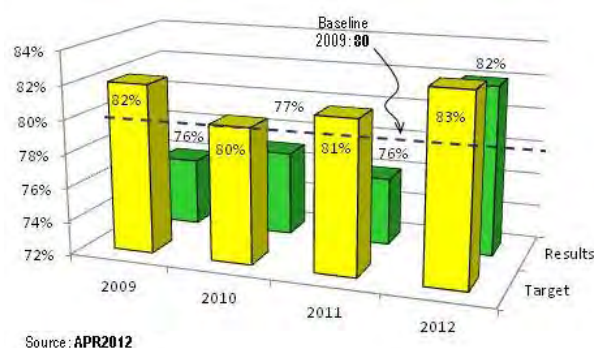


Figure II.30: % of enrolled primary students attending

expectations. Officials at *Save the Children* have indicated they noticed, over the years, a reduction in the population of children of pre-school age in the bateys intervened; which might explain the backward movement in the value of this indicator. Now, this trend would have been a cogent reason to revise the annual targets, as it was admittedly done in the case of other indicators; yet, it was not done in this case.

2.28 Early outcomes appear also somehow disappointing in relation to the proportion of *enrolled primary school students who actually attended classes* in the bateys intervened; indicator for which also a baseline value was made available. That percentage was consistently falling below target -and below the baseline performance as well- in all the first three years of the Project. (See figure II.30). It too failed to achieve the target on the final year, but at least then it registered two percentage points above the baseline value of 80% attendance, prior the initiation of the Project. Similar results are seen in the enrollment of children and youth in extra-curricular programs, as depicted in figure II.31. Results on this indicator fell below annual target as well, on three of the four years of the Project; performing better than expected only on 2009.

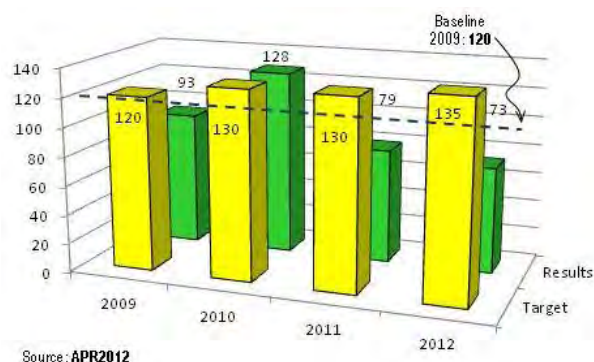


Figure II.29: # of children enrolled in pre-school

To begin with, the Project did not achieve the stipulated annual target for this indicator on any year of execution. But, most importantly, at the close of the operation results ended up at 73 children enrolled -just above half of what was expected for that final year- and a level that represented a 39 points actual drop in pre-school enrollment, if compared to the levels observed before the start of the Project, as is illustrated in figure II.29. Therefore, the observed change with which the Project can be “associated” has run contrary to

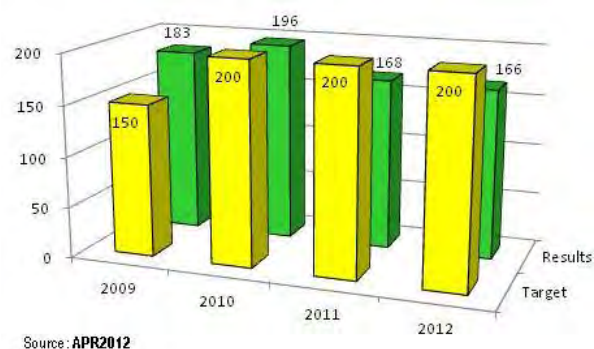


Figure II.31: # of 6-14 year olds enrolled on extra-curricular programs

2.29 Finally, on the # of youth and adults participating in vocational or informal education, results exceeded targets in 2009 and 2010, but fell below expectations on the 2011 and 2012. (See figure II.32) None of these two last indicators included base-line measurement; so no estimation of changes can be established with respect to the situation prior to the initiation of the Project.

(iii) Sanitation

2.30 In what concerns the improvement of sanitation conditions, the *number of families with improved access to clean drinking water* in the intervened bateys is another terminal output indicator that can be

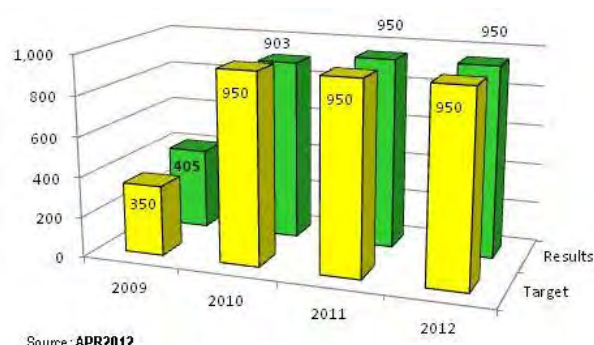


Figure II.33: # of families with improved access to clean drinking water

considered very close to the frontier with very early development outcomes of the Project, because the first action toward generally better health conditions is very possibly just a small step away: i.e. the act of habitually drinking pure water. According to the data summarized in figure II.33, the number of families who were provided by the Project with such improved access -through the distribution to families of water filters for use at home- corresponded to, or surpassed the annually planned target in all years of project execution, with the only exception of year 2010; year in which actual results were below the stipulated target. Since no base-line value was provided for this indicator, no change can be measured either, with respect to the situation prior to the Project.

b. Terminal outcomes

2.31 More advanced development effects induced by the BDCP are measured with the seven indicators included in our working *LogFrame* matrix as corresponding to terminal outcomes of the Project. These few measurements are clearly separate from project outputs and at least approximate initial changes in behaviors that are discernible in the batey population, concerning habits of health, wellness and nutrition, as well as education results. Four of these seven indicators do provide baseline values, therefore affording also a measure of changes in such behaviors and habits whose start can be time-associated with the workings of the intervention. The first of these positive outcomes is represented by the *percentage of care givers reporting three key child health & nutrition practices*. Certainly significant results are reported by Management in terms of actual observation of those positive child health and nutrition practices by care givers in the intervened communities. Figure II.34 illustrates how the targets of % of care givers reporting such

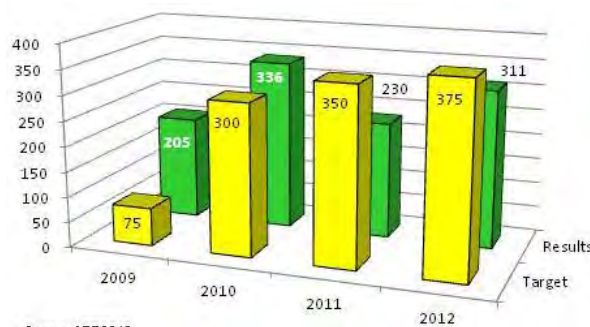


Figure II.32: # of youth & adults participating in vocational or informal education

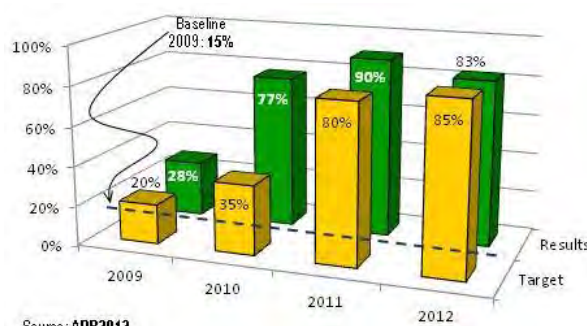


Figure II.34: % of care givers reporting 3 key child health & nutrition practices

practices were exceeded on all years of the operation, and ended up registering an increase from 15% of individuals reporting such practices before the start of the BCDP to 83% of them doing so at the end of the Project: a very significant change that has been concomitant with the execution of the Project.

2.32 Annual targets are also reported to have been exceeded on the *proportion of babies under 6 months exclusively breastfed in past 24 hours* at the point of measurement. This performance is illustrated in figure II.35, with actual results in fact surpassing annually set targets for all years of the operation, except for the starting year of 2009. For this year no target was set, or performance was measured. Also, since there is no baseline value for this indicator, no cogent argument can be made about a measurable difference associated with the Project on this particular respect.

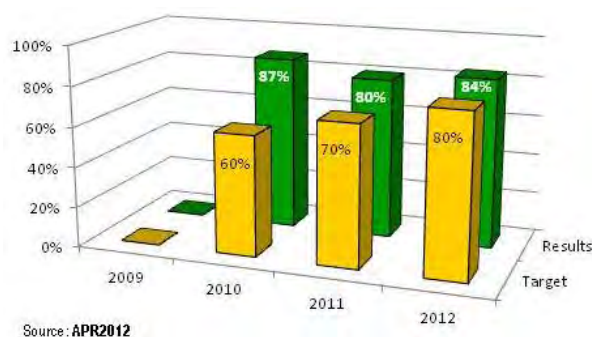


Figure II.35: % of infants under 6 months exclusively breast-fed in the past 24h

2.33 Other strong indication of positive advanced development outcomes likely associated with the Project, concerned the percentage of *women ages 15-49 reporting 2 key reproductive health practices*. Figure II.36 depicts how results also surpassed all annual targets, and the Project appears to have contributed to important women health practices, multiplying by a factor of two the percentage of females reporting such practices from the baseline point to the project-end point. This very significant change in practices, along with the already discussed early outcome concerning women receiving PAP smear tests in significant numbers, tends to ratify important transformations already occurring in women wellness and health habits, and that are probably associated with the workings of the BCDP intervention in the targeted batey communities.

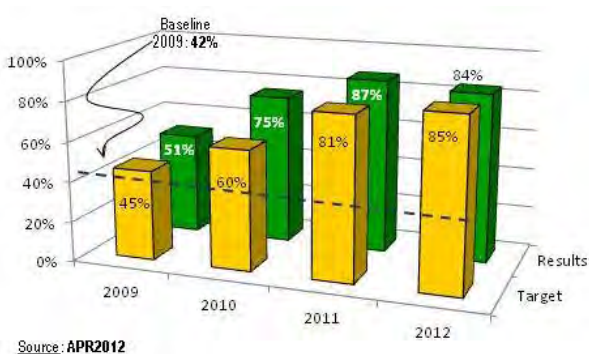


Figure II.36: % of women ages 15-49 reporting 2 key reproductive health practices

2.34 Finally concerning terminal outcomes on health, Management reports performance in the communities of phase I with the following proxy indicator on adult wellness practices: (i) *% of adults identifying two STI and HIV/AIDS prevention practices*; and with the following indicators of children hygiene practices measured by: (ii) *% of primary school children washing hands after using school latrine/bathroom*; and (iii) *% of primary school children washing hands before eating at school*. A more advanced health outcome is also shown with the indicator: (iv) *% of children under 5 with diarrhea in the last two weeks*. Performance on the first of these indicators is depicted in figure II.37. As can be seen, annual results are mixed. After targets were surpassed the first two years of the Project, results

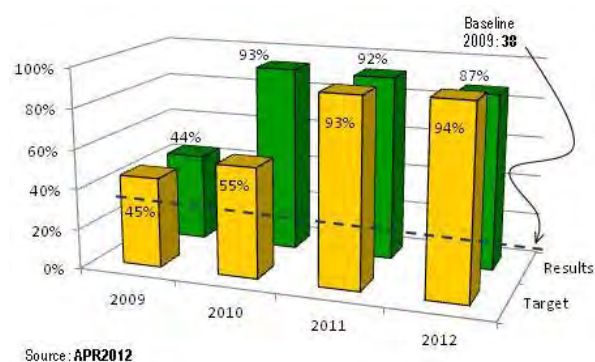
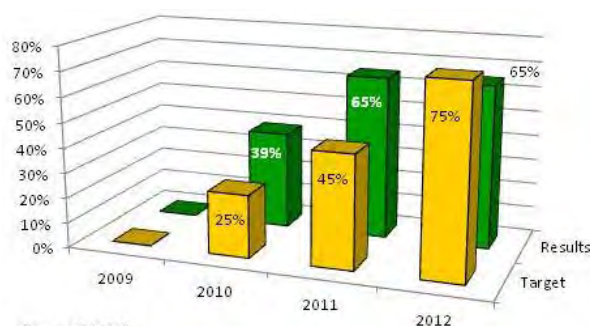


Figure II.37: % of adults identifying two STI and HIV/AIDS prevention practices

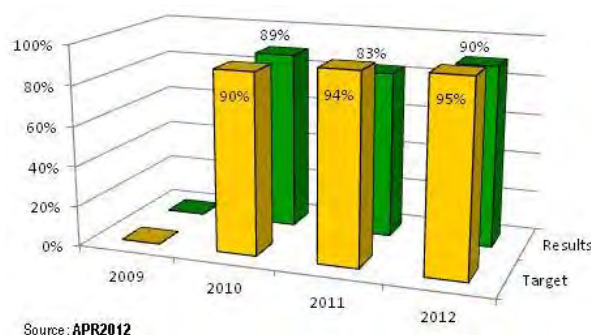
actually fell below targets in the last two; with an accumulated final 87% of adults identifying two STI and HIV/AIDS prevention practices by the end of 2012. This represented a step back from both the target set and the results achieved in 2011; nevertheless, that final percentage represents a 49 percentage point increase from the baseline value. Once again, this is a significant change that can be associated with the execution of the Project.

- 2.35 A less shiny performance is reported concerning the indicators about children's hygiene behaviors at the school. Figure II.38 shows that, with no available target or measured results for 2009 in terms of children washing their hands after going to the toilette at school, results exceeded targets in the middle years of 2010 and 2011. However, on the final year 2012, results fell below the annual target, and only kept on a par with the results achieved in 2011.



Source: APR2012

Figure II.38: % of primary school children washing hands after using school latrine/bathroom



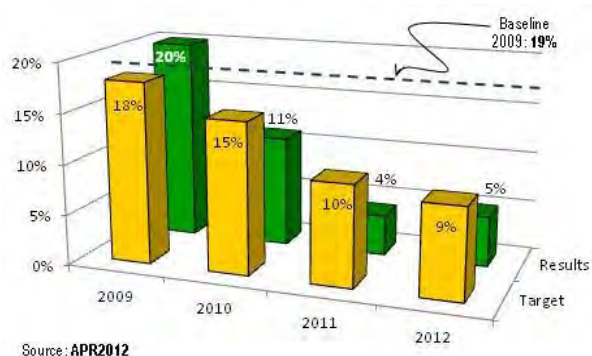
Source: APR2012

Figure II.39: % of primary school children washing hands before eating at school

Figure II.39 shows the same lack of information for year 2009 on children's habit of washing hands before eating at school, as with the previous indicator. Also in the following years, results consistently fell below annual targets, with a final cumulative percentage of 90%, of all kids observing the mentioned practice; which was only the target for the starting year of measurement. Admittedly, these results represent a quite high incidence of a positive hygiene behavior in children; unfortunately, in the case of this indicator as with the previous one, the lack of baseline values

precludes us from having any idea of differences with the ex ante situation that may have been induced by the BCDP concerning these important preventative behavior.

- 2.36 Despite the above, the even more advanced project outcome, observed in the incidence of diarrhea in small children, is very likely the consequence of emerging hygiene habits and sanitation conditions in the intervened bateys, and constitutes a result in which a pronounced change can clearly be associated with the Project. Indeed the *percentage of children under 5 with diarrhea in the two weeks* prior to measurement, which was supposed to decrease, registered at 20% during the first year of the Project: two percentage points above the target for the year, and one percentage point above the measured value of the indicator before the execution of the Project. However, in all years after 2009 annual targets were consistently exceeded by the Project; which, at the end, managed to reduce the incidence of diarrhea as defined in the indicator from 19% to only 5% of all small children of the treatment group; as can be seen in figure II.40.



Source: APR2012

Figure II.40: % of children under-5 w/diarrhea in last 2 weeks

2.37 The last of the terminal outcome indicators, in our working LogFrame matrix, corresponds to the area of Education, and is the *% of students enrolled in 4th grade that pass 4th grade*. Performance on this indicator is illustrated in figure II.41 which depicts good results as compared with set out targets for all years, except for 2011. It is particularly noteworthy that the proportion of 4th grade students passing that grade on 2009 stood at a relatively strong 84% on a year for which no target had been established by the Project. That passing rate was identical to the baseline value of the indicator, against which to compare performance in the rest of project years. By the end of the Project the passing rate was reportedly improved by 11 additional percentage points among 4th grade students in the intervened bateys.⁶

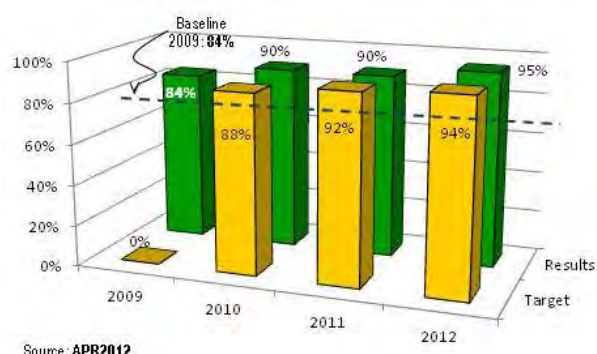


Figure II.41: % of students enrolled in 4th grade that pass 4th grade

2.38 Finally, targets and results for project outputs and outcomes in the new bateys of phase II of the Project, as reported in APR2012, are presented in table II.1⁷

#	Indicator	SECOND PHASE (2011 - 2013) – 4 new communities and 2 new Areas												
		Baseline 4 new communities 2010 (Nov)	Target 2011 (sept)	Result 2011 (sept)	Target 2012 (sept)	Result 2012 (sept)	Target VB 2011 (Sept)	Result VB 2011 (sept)	Target VB 2012 (sept)	Result VB 2012 (sept)	Target Bahor. 2011 (sept)	Result Bahor. 2011 (Sept)	Target Bahor. 2012 (Sept)	Result Bahor. 2012 (Sept)
1.1	# of people benefitting from USG supported health services	0	700	775	775**	775	3,500	2,340	4,500	9,604	1,600	1,845	1,845	1,845
1.2	% of children under five with diarrhea in the last two weeks	27%	20%	9%	16%	4%								
1.3	% children under 12 months of age who have received DPT3	50%	65%	75%	80%	88%								
1.4	% of children under 5 years of age who received vitamin A from USG supported programs			80%	85%	86%								
1.5	% of children under 10 years of age who are dewormed by USG-funded programs			90%	90%	89%								
1.6	# of Community Health Promoters trained in: • Maternal/newborn health • Child health and nutrition • Reproductive health/family planning • TB management	0 0 0 0	8 8 8 8	8 8 8 8	8 8 8 8	8 8 8 8	4 4 4 4	4 4 4 4	4 4 4 4	8 8 8 8	4 4 4 4	4 4 4 4	4 4 4 4	4 4 4 4
1.7	# of Community Health Promoters equipped	0	8	8	8	8	4	4	4	4	4	4	4	4
1.8	# of bateyes with functional MISPAS reference systems	-	4	4	4	4								

Table II.1: Annual targets and achievements in the bateys of BCDP phase II. Source:

⁶ Information on the important indicator: *% of students in 1-4 grade of primary schools of who read at grade level* was not disclosed by Management in the APR2012 operational report.

⁷ The information is quoted *verbatim* from the tables in the executor's report

#	Indicator	SECOND PHASE (2011 - 2013) – 4 new communities and 2 new Areas												
		Baseline 4 new communities 2010 (Nov)	Target 2011 (sept)	Result 2011 (sept)	Target 2012 (sept)	Result 2012 (sept)	Target VB 2011 (Sept)	Result VB 2011 (sept)	Target VB 2012 (sept)	Result VB 2012 (sept)	Target Bahor. 2011 (sept)	Result Bahor. 2011 (Sept)	Target Bahor. 2012 (Sept)	Result Bahor. 2012 (Sept)
1.9	% of care-takers who report at least three key positive child health and nutrition practices	26%	35%	79%	60%	89%								
1.10	% infants under 6 months of age exclusively breastfed within the past 24 hours	22%	30%	56%	40%	50%								
1.11	% women between the ages of 15 and 49 receiving at least one PAP smear test in the past 12 months	37%	50%	64%	60%	81%								
1.12	% of women between the ages of 15 and 49 who can report at least two key reproductive health practices	36%	50%	96%	70%	94%								
1.13	% of heads of households who identify at least key two STI and HIV/AIDS prevention practices	34%	68%	81%	80%	99%								
1.14	# of adults tested for TB	0	10	24	15	16								
1.15	% school-children who wash their hands: - After using the school latrine/bathroom - Before eating the school breakfast, snack, lunch	25% 74%	30% 80%	47% 84%	60% 90%	65% 90%								
2.1	# classrooms rehabilitated and equipped	0	10	12	10	13								
2.2	# of classrooms with improved didactic materials	0	10	12	10	13								
2.3	# of teachers from participating batey schools trained in: - Literacy - Math - Computer use - School Health and Nutrition - Other pedagogical improvements/practices	0 0 0 0 0	4 4 13 11 11	6 0 0 0 12	4 4 13 11 11	8 0 4 0 12								
2.4	# of administrators and education officials trained	0	7	4	7	4								
2.5	% of enrolled primary school students attending	74% (February 2011)	80%	76%	82%	85%								
2.6	% of students enrolled in 4 th grade who pass 4 th grade	88% (measured in June 2011)	90%	96%	92%	95%								
2.7	% of students in single grade primary schools in 1-4 grade who read at grade level	TBD	TBD	TBD	TBD	---								
2.8	# of children enrolled in pre-school	36	40	18	45	29								
2.9	# of children and youth (6-14 years old) enrolled in USG supported extra-curricular programs	0	60 (1)	116	75 (1)	102								
2.10	# of youth and adults participating in USG-supported vocational or informal education programs	0	135 (1)	45	250 (1)	132								

Table II.1 (continued): Annual targets and achievements in the bateys of BCDP, phase II.

#	Indicator	SECOND PHASE (2011 - 2013) – 4 new communities and 2 new Areas												
		Baseline 4 new communities 2010 (Nov)	Target 2011 (sept)	Result 2011 (sept)	Target 2012 (sept)	Result 2012 (sept)	Target VB 2011 (Sept)	Result VB 2011 (sept)	Target VB 2012 (sept)	Result VB 2012 (sept)	Target Bahor. 2011 (sept)	Result Bahor. 2011 (Sept)	Target Bahor. 2012 (Sept)	Result Bahor. 2012 (Sept)
3.1	# of families with access to improved sanitation services	0	50	0	100	203					25	25	50	50
3.2	# of families benefitting from improved community water systems	0	175	0	350	372								
3.3	# of families with improved access to clean drinking water	0	150	16	300	76	275	0	350	10	150	4	250	12
3.4	# families with improved housing	0	20	0	50	176								
3.5	# of bateys with access to an emergency-safe structure	0	1	0	2	3								
3.6	# of bateys with updated emergency plan	0	4	4	4	4								
3.7	# of batey schools with updated emergency plan	0	4	4	4	4								
3.8	# of residents receiving support in income generation activities	0	50	26	130	83								

Table II.1 (continued): Annual targets and achievements in the bateys of BCDP, phase II.

III. Findings 2: Attributable results

- 3.1 Having registered above the Project results as reported by executors, in the present section the researchers proceed to discuss independent empirical evidence on such results, to the extent that it is available based on the field data gathered and compiled by both the separate studies done previously at the Project's *baseline* and *project-end* points, and the field study done presently on the control groups. Whenever it was possible, or warranted, results and comparisons done by the present evaluation have been subjected to statistical significance test (χ^2) and related parameters (**p**; **w** and **power**), so that observed gaps between treatment and control groups can be imputed to the Project (*attribution criterion*) following the quasi-experimental study design referred to in paragraph 1.10 of the present report. Also, the discussion that follows presents the testing of the BCDP *development hypothesis* (expected results chain) as reconstructed in the *Note on the Evaluability of the Project* (paragraphs 21 and 22) and in the *Introduction* (paragraphs 1.3 and 1.4) of the present report. Therefore, the presentation of empirical findings in the following sections is ordered and structured around the said hypothesis; according to which: the submission by the Project of a number of deliverables (*development outputs*, both intermediate and terminal) is expected to have promoted a number of transformations in the bateys' living conditions (early or *intermediate development outcomes*) which, in turn, is expected to have induced positive initial behavioral changes in batey residents (*terminal development outcomes*).
- 3.2 Given the information limitations discussed in the *Note on the Evaluability of the Project*, the full scope of the reconstructed evaluation framework and quasi-experimental design could be applied, and taken advantage of, only on the data gathered from the bateys in the *San Pedro de Macorís–Hato Mayor* region -phases I and II-; and, even there, with frequent exceptions. This is the reason why the main discussions that follow refer to that particular region; and why the other, more disjoint findings are discussed separately. Independently gathered data are generally preferred; yet, frequently on particular comparisons, data from the Executors' final operational report (APR2012) are used in order to make up for the lack of data from prior field surveys. Finally, and because of the characteristics of the majority of indicators identified in the *Final List* of metrics for the BCDP, a discussion of the parameters for

statistical significance provided by the χ^2 test will be included in the text only when warranted by the nature of the indicator in question, in accordance with the points discussed on paragraphs 16 through 19 of the *Note on the Evaluability of the Project*.⁸

A. Quantitative perspective: SPM-Hato Mayor bateys, phases I and II

1. Project outputs

3.3 The chief project outputs to be evaluated herein refer to enhanced infrastructure and services that the operation provided the intervened bateys, mainly in the areas of Health, Education and Sanitation.

a. Health

(i) Attention & treatment

3.4 During project execution direct health attention and treatment services were provided to the intervened batey population, both in Phase I and 2; so a clearly increased number of individuals benefited from those services. Figure III.1 illustrates the Project's performance in terms of this deliverable, from the zero coverage value at the base-line. As it was expected, starting also from no services received, the not-intervened control group remained flat at zero in this indicator. For this reason, no application of results from the *Chi-square* test is warranted in the case of the present indicator.

3.5 It must be pointed out, however, that there is some ambiguity in the result presented herein due to the fact that the Executor, having defined the indicator as presented, for no clear reason changed afterward the meaning of the metrics used. Indeed, researchers found that, although the name of the indicator only refers to people benefitting from health related services, the Executor's final *Performance Report* (APR2012) defines it as referring to "one or more activities of the project, especially activities of the *health* and *education* components". For this reason, it must be noted that the measure actually presented here is far more extensive than required by the indicator.

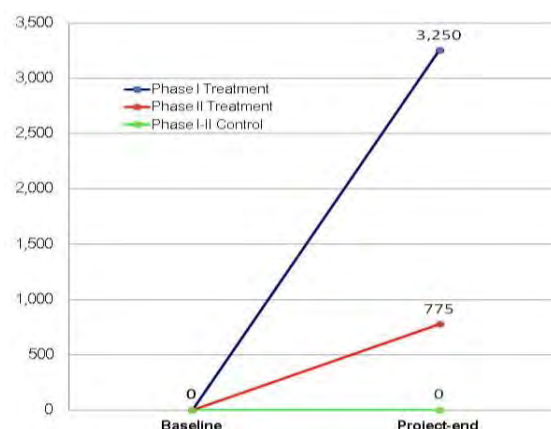


Figure III.1: # of people benefitting from health services.

3.6 The perception of the interviewees on the Final study validates the APR2012 claim that the total population of the phase I bateys (3,250) benefited from one or more activities of the Project, in the broader sense indicated above. For the Phase II bateys, the report only points out that the Project reached its target of 775 persons, but the mentioned perception on the Project-end study confirms, in fact, that 100% of the target population of this Phase II in SPM-Hato Mayor benefitted from the Project.

⁸ There it was shown that when the base-line and end-line values are zero for an output, the χ^2 test does not add value to the analysis. Now, for the cases with a non-zero value, such test is applied in this chapter.

3.7 Also, measures of other output indicators confirm in a more specific way the provision of important health services such as children immunization and de-worming. Figure III.2 depicts project performance in terms of infants receiving DPT3 in the treatment and control bateys. The Baseline study show the % of babies under-one year receiving DPT3 to be near the 50% on both control and treatment groups in the phase II; while in the treatment group of phase I the corresponding value was 13%. Before continuing the analysis, it must be noted that this is an exceptional *output* indicator, in that: (i) it is in the frontier with *outcomes*; and (ii) there is a base-line measuring the output delivered at the start, obviously not by the Project. Thus, a non-zero value was registered for the control group at the base line, revealing the salient fact that before the BCDP roughly half of all babies did receive DPT3 in all communities of phase II. Now, by the end of the Project, while the proportion of infants targeted in both treatment groups considerably increased to around 90% of all babies, in the Control Group the percentage actually descended dramatically to only 10%.

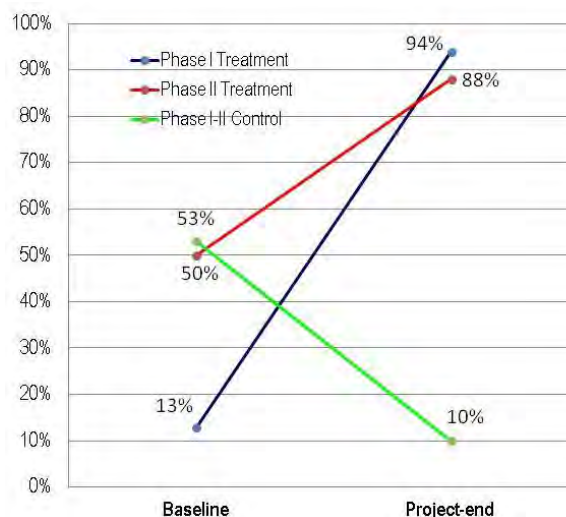


Figure III.2: % of infants under 12 month receiving DPT3

3.8 The above suggests that while the Project itself was very successful in covering the infant population in the intervened bateys, such success appears unfortunately correlated with a marked deterioration in the services to infants in bateys not intervened by BCDP, who used to receive DPT3 from other sources, but, for some reason, stopped receiving it as before. The statistical significance of these differences is high, as the *Chi-square*₂ test yielded a value of 87.718, with $p = .000$. The effect size is large too, with a w value of .53 and the **power** of the analysis also yields a very high value of .99. These parameters substantiate a high confidence in the fact that the important gaps observed between Treatment and Control groups are not due to chance and, therefore, are pretty well correlated to the workings of BCDP; including the unfortunate possibility that government vaccination programs in bateys, nearby the Project targets, had actually diminished their incidence.⁹

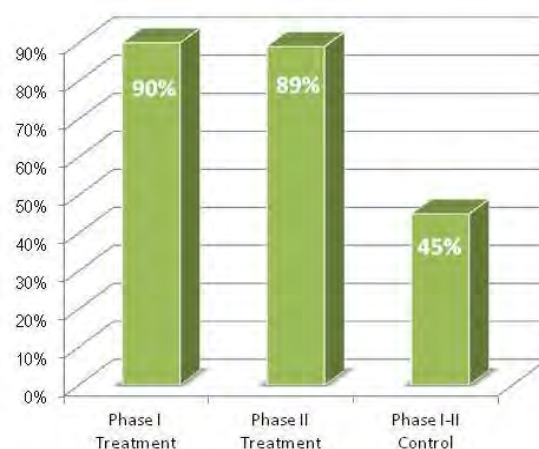


Figure III.3: % of children under 10 being de-wormed

3.9 Measurements also attest to a high level of project delivery in terms of de-worming children under-ten years of age, as shown in figure III.3. However for this indicator, which was added in 2011 for the extension of the Project, no baseline study was performed; which, in turn, limits the present report to

⁹ In discussions with executors, the possibility has been raised that, given a fixed -and probably insufficient- supply of vaccines provided by the Government to the whole area, a more efficient use of them by the Project in the Treatment bateys may have caused the decrease of vaccinations in the not-intervened bateys. Now, however plausible an explanation might be for the conspicuous result, this hypothesis was not independently confirmed.

submit results only from the Project-end study, with no comparison possible to the situation before the Project. While in the treatment groups of both phases we see close to 90% of children being de-wormed, the control group only registers 45%. The figure shows the important difference found between treatment and control groups relating to this indicator; a difference that turned out also to be statistically significant, with a χ^2 of 17.688 and $p = .000$. The effect size is medium, with w of .33 and the **power** of the analysis yields a very high value of .99.

- 3.10 Finally, field measurements also confirm the Project's delivery in terms of the establishment of official health Reference Systems of the government relevant ministry (MISPAS). As the indicator of children de-wormed, the indicator of *# of bateys with functional MISPAS reference systems* was added in 2011 for the extension of the Project and not included in the baseline studies performed in 2009. That is why we, also here, limit our report to comparing results of the Project-end studies only. These results are shown in figure III.4. The treatment values of 8 MISPAS Reference Systems installed and functioning in phase I and 4 in phase II were supplied by the APR2012; while the value for the control group was found in our evaluation survey, where 85% of the interviewees in the Batey *Alejandro Bass* mentioned a MISPAS reference system in the community where they could go by foot. However, while the Control Group has a much lesser number of MISPAS Reference Systems functioning, this time the difference among the groups did not reach statistical significance, with a χ^2 value of 4.38, and $p = .803$).

(ii) Wellness promotion

- 3.11 The Project also delivered in the effort to promote the adoption of good health practices and habits among batey residents. This was done specifically through the training and equipment of community health promoters in the intervened bateys, as well as preventative testing on individuals. Concerning the community health promoters, the original metrics designated separate indicators for training and for equipping health promoters; yet, for the sake of simplicity, we treat these two indicators together in the analysis (figure III.5) because the values are the same for both of them, making redundant two separate analyses. In this case, the Treatment information was supplied by the APR2012. The training of promoters included the subjects of (i) Maternal/newborn health; (ii) Child health and nutrition; (iii) Reproductive health/family planning; y (iv) TB management.

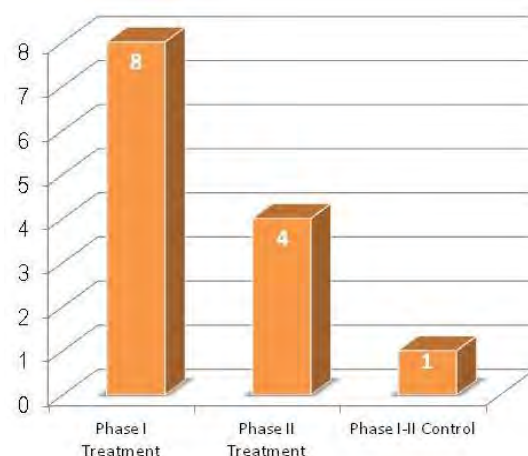


Figure III.4: # of bateys with MISPAS reference systems

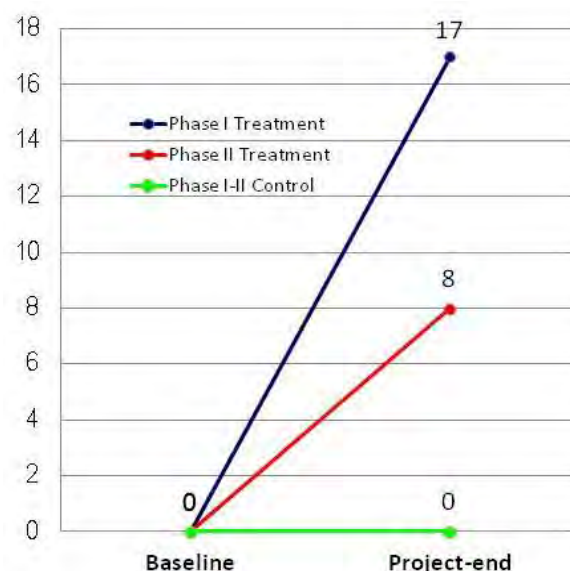


Figure III.5: # of health promoters trained & equipped

3.12 Results on the indicator for preventative testing of TB appear to show a relatively poor project performance as compared to that of other possible testing sources. Concerning the output of individuals tested for TB, this indicator was not included in the baseline databases for the phase I and II treatment groups or the Control Group. The APR2012 reports a 0 value for the phase II treatment group baseline. Thus, once again, we cannot compare baseline values with end-project values, nor can we apply our full Treatment-Control comparison, for this particular case. Instead we are limited to only report the results of the Project-end studies, as shown in figure III.6.

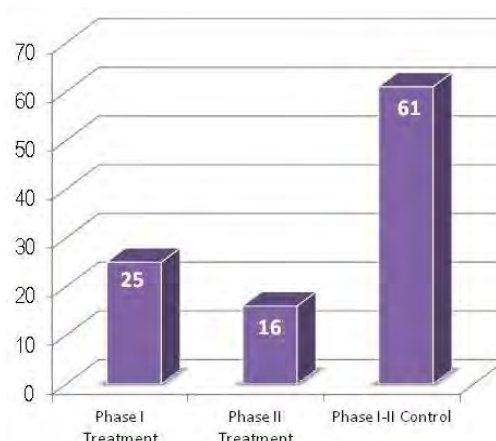


Figure III.6: # of adults tested for TB

3.13 The researchers identified a practice, during the Project's execution, that is clearly concomitant with this particular result; namely: that the Project only tested people who were TB symptomatic and, as result, in the treatment group only 25 individuals turned out to be so in phase I, as well as 16 individuals in the phase II. From those tested, the Executors reported that all positive cases were eventually medically treated. Now, the intriguing fact is that in the control group survey the interviewees reported a very high number of 61 persons as having being tested for TB, one third of them coming from the *Alejandro Bass* batey.

3.14 Now, apart from the possible role of the mentioned difference in testing protocols, this unexpected result is also possibly associated with the correct operation of the MISPAS reference system that the researchers detected precisely on the community of *Alejandro Bass*. If this was the case, -although no independent confirming evidence was discovered- it would seem like the government intervention actually outperformed the Project on this particular respect, in communities not intervened by the BCDP.

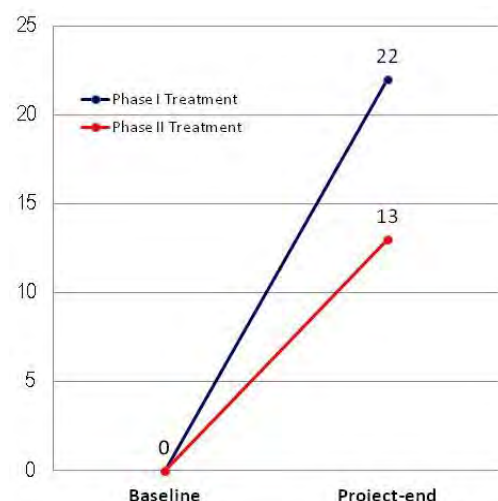


Figure III.7: # of classrooms refurbished & with improved didactic materials

b. Education

(i) Refurbishing of school infrastructure

3.15 The original metrics design included separate indicators for *# of classrooms rehabilitated and equipped* and *# of classrooms with improved didactic material*. As it was the case with the indicators on community health promoters, here we analyzed these two indicators together because the values are the same for both of them (See figure III.7). Being the number of classrooms actually refurbished and equipped not a matter of opinion, questions about that were not included in our evaluation survey, as it was not appropriate to ask for an "estimate" on the matter from the interviewees. Instead, we use the information on results in the treatment group supplied by the executor report APR2012. Therefore, we only present the information relevant to the treatment groups from the baseline and Project-end studies in phases I and II.

(ii) Enhancement of teacher competencies & tools

3.16 Arguably, the indicator on classrooms provided with improved didactic material gives already a measure of how the teachers were given enhanced tools for pedagogical purposes. But teachers also received training to improve their competencies on several important subject matters. This effort at training also included school officials and administrators on subjects relevant to their respective functions.

3.17 One important measure of the Project's performance in this respect is the *number of teachers from participating batey schools trained in Literacy* and its different components. Results on this indicator, including the values concerning the different components of such training, were supplied by the APR2012 from the baseline and Project-end studies on phases I and II, and are illustrated in figure III.8.

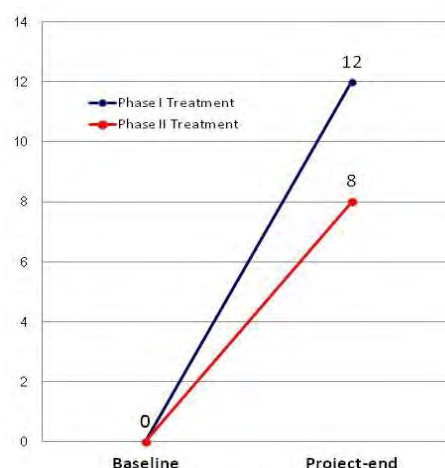


Figure III.8: # of teachers trained in literacy

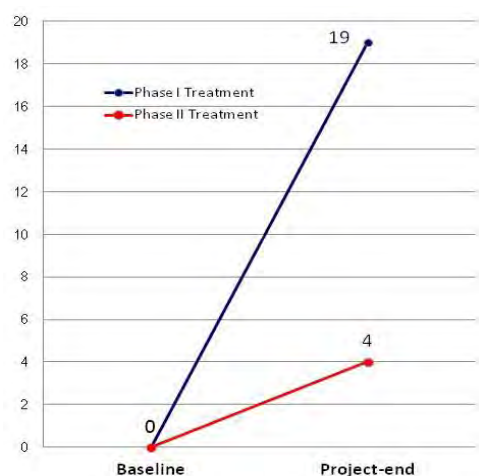


Figure III.9 # of teachers trained in Computers

figure III.9.

3.18 In what concerns the number of teachers from participating batey schools trained in School Health and Nutrition, project outputs appeared only submitted in the bateys of phase I. The only value reported in the APR2012 was 12 teachers from participating batey schools in the treatment group of phase I; and, therefore, results for phase II are shown to be zero, as illustrated in figure III.10.

3.19 Beside the subject matters already reported on, batey school teachers were also trained by the Project in other pedagogical improvements and practices, so that their teaching competencies would be enhanced. In this respect it can be reported that the number of teachers benefitted reached significant numbers in the intervened batey schools.

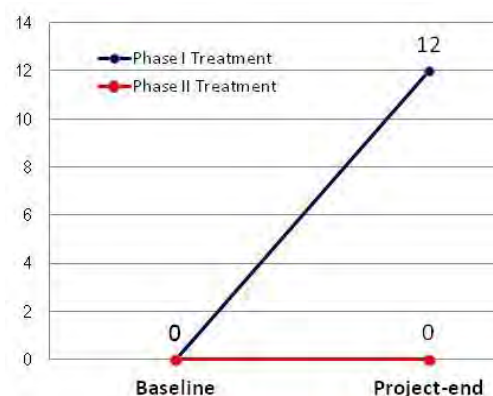


Figure III.10: # of teachers trained in School Health and Nutrition

3.20 The training of teachers in other pedagogical improvements or practices in the participating batey schools included the subject matters of: a) positive discipline; b) multiple intelligence; c) planning, and d) participatory didactic methodologies. The number of teachers whose competencies on those subject matters were enhanced by the Project reached 13 in the phase I treatment group and 12 in the phase II treatment group, as can be seen in the figure III.11.

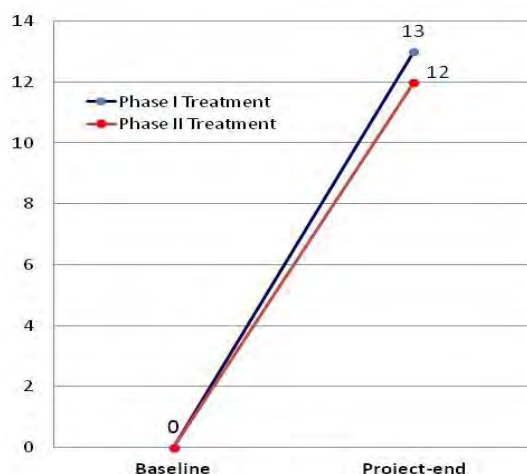


Figure III.11: # of teachers trained in other practices

3.21 Also, as we mentioned before in the present report, administrators and education officials of the intervened batey schools received special training as well (See figure III.12). Specifically, their competencies were enhanced on how to prepare a *management plan* for each school, based on the SWOT analysis: a subject matter relevant to their functions in the education system.

3.22 The data available reveals that the number of administrators and education officials trained in this subject matter was 8 for the phase I treatment group and 4 for the phase II treatment group. As in all the above indicators these *output* results in the targeted bateys were not the subject of the field survey done by the present researchers -because the latter was directed to the not intervened bateys- and so the present analysis is based on the information supplied by the APR2012 report.

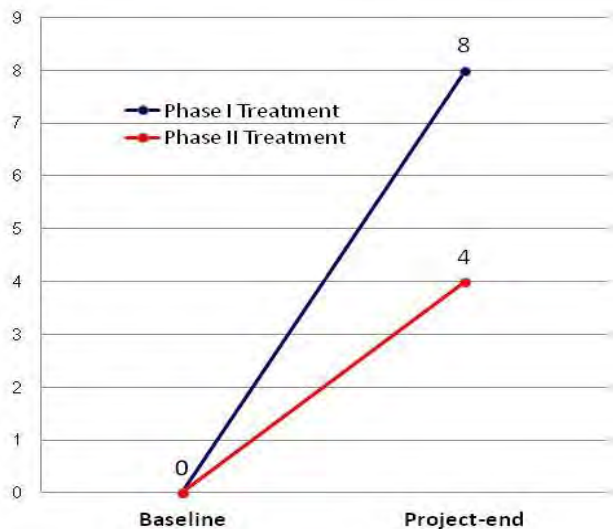


Figure III.12: # of administrators, education officials trained.

c. Sanitation

(i) Revamping of sanitary infrastructure

3.23 The Project delivered significant products in terms of repairing and improving the intervened bateys' sanitation infrastructure, which was to be measured by the *number of families with access to improved sanitation services*.

3.24 Contrary to the situation concerning the refurbishing of school infrastructure, discussed in the preceding section, improvements in sanitation infrastructure could in part be corroborated during our evaluation surveys of families, especially in what concerns the presence or absence of latrines. Therefore, the researchers included in the surveys questions already posed in previous studies to validate the results obtained on the above mentioned indicator in the control bateys, for comparison with targeted ones. These validation measures and the corresponding findings by the field research are discussed below, to complement the findings as measured by the main indicator on improvements of sanitation infrastructure from the baseline to the Project-end point.

3.25 The data on families with access to improved sanitation services as a result of the BCDP was supplied by the APR2012 report, based on the latrines installed. As illustrated in figure III.13, the available information indicates that the number of latrines constructed was 300 in the phase I and 203 in the Phase II.

3.26 As indicated above, in addition to this information provided by the Executors report, in the final control survey the researchers repeated questions posed in previous surveys that provide a means to validate the above results, based on direct testimony on the particular output of latrines installed. One validating question on latrines focused on whether they functioned or not. On

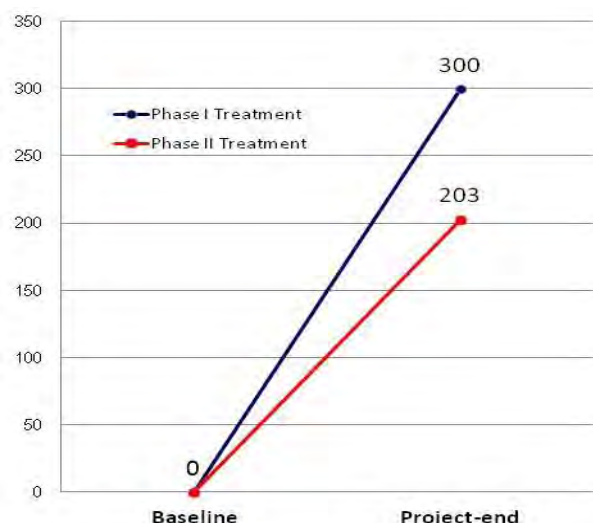


Figure III.13: # of families with access to improved sanitation services.

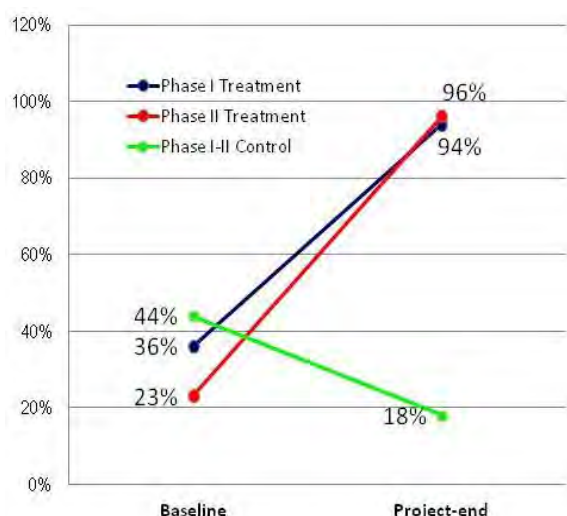


Figure III.14: % of latrines functioning always, almost always.

this respect it was found that, at the end, the % of latrines functioning always or almost always was quite high in the targeted bateys (See figure III.14) Since on this particular metrics there was available baseline data on both control and treatment groups, it is possible to measure the corresponding trend and differences between the groups. Notice how, starting from a lower value than that in the control group, the indicator notably increases for the intervened population. At the same time it actually decreases in the control group, from the higher value it had to begin with.

3.27 As is shown in the figure, the baseline positive report about the functionality of the latrines ranged from 23% to 44%, but the Project-end studies of the treatment group yielded a percentage of 94% for those families in phase I, and of 96% for those in phase II. Most importantly, in our survey of the control group this positive report of functionality descends to 18%.

3.28 The clear differences in functionality of the sanitation services are very likely due to the workings of the Project. The tests yielded statistical significance, with a Chi-square₂ = 52.044, **p** = .000; a medium size effect (**w** = .43) and a very high **power** of .99. This indicates that a substantial difference was made by the BCDP for the intervened population, in terms of this particular development output.

3.29 Similar results have been observed when measuring the percentage of existing private latrines, and the same positive effect of the Project can be observed on this respect. Starting from percentages ranging from 14% to 36%, the Project-end studies on the treatment group reported that 77% of families had private latrines (for the exclusive use of the household) in phase I, and 80% in phase II. In the control group, on the other hand, the percentage of private latrines actually fell from 36% in the baseline to

25% in the Project-end study. (See figure III.15) These differences showed statistical significance ($\chi^2 = 34.486$, $p = .000$) a medium effect size ($w = .39$) and a very high power (.99).

- 3.30 The measurement of actual use of latrines by family members yielded results along the same line. Here again, a general positive effect of the Project can be seen in figure III.16 on the percentage of families where every member of the household uses the latrine in the Treatment and Control Groups. Once more, the baseline studies show relatively low percentages ranging from 25% to 42%, while the Project-end studies for the treatment groups report a percentage of 95% of households in phase I and

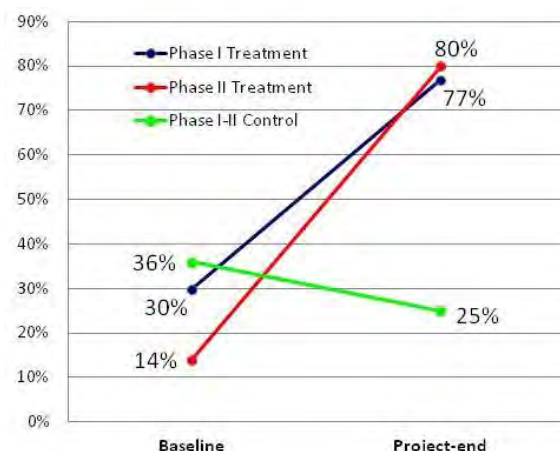


Figure III.15: % of families with private latrines.

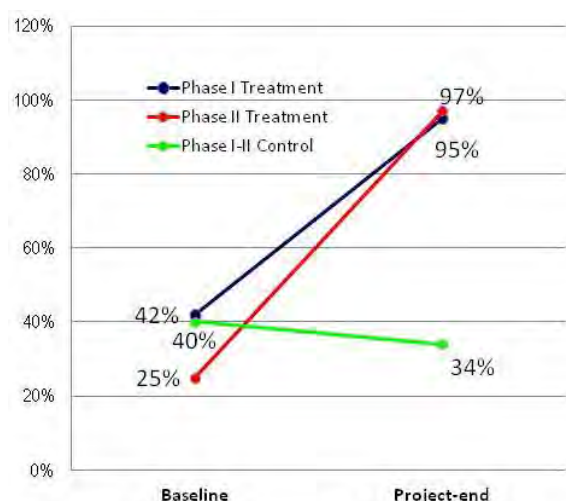


Figure III.16 % of households where every member uses the latrine.

97% of families in phase I, where every member of the household uses the latrine. The percentage in the control group falls from 42% in the baseline to 34% in the Project-end study. These differences are statistically significant. The analysis yielded a χ^2 of 24.024, with $p = .000$; a medium effect size ($w = .27$); and a very high power (.99). Finally the measurement of trends in the access to latrines through the years clearly confirms that the Project delivered on this respect. In addition to the above discussed output results, we compared the percentages of households who in the Project-end study reported having a latrine on the intervened communities by years, and this information is shown in the figure III.17.

- 3.31 Having had no change in the possession of latrines in the years prior to the start of the BCDP in 2009, families began to report increasing availability of latrines thereafter. In the figure we can observe that the percentages of households owning a latrine start to ascend by 2009 in phase I treatment group, precisely by the beginning of the Project, until the proportion reached 87% of all households by the end. As phase II of the Project started in 2010, that very year the possession of a latrine begin to ascend until, by the end of the project, it reached a 97% of all households intervened in that second phase.

- 3.32 The yearly differences for the phase I treatment group are statistically significant. $\chi^2_6 = 24.965$, $p = .000$. The effect size is medium ($w = .25$) and the power of the analysis is very high (.98). We also found



Figure III.17 % of households reporting a latrine by year.

statistical significance for the yearly differences in the Phase II Treatment Group ($\chi^2_6 = 38.586$, $p = .000$); with a medium effect size ($w = .31$) and perfect power (1).

(ii) Upgrading of water systems

3.33 Finally in the area of sanitation, another important development output delivered by the BCDP in the targeted batey communities was the improvement of water systems for collective use. The systems were not only upgraded by the Project in what concerned their physical infrastructure, but also in the very important sense that a community self-management mechanism was promoted and implanted by the Project, for the operation and maintenance of the systems; promoting, therefore, their future sustainability. The metrics designated to measure this project output was the *number of families benefitting from the improved community water systems*.

3.34 The values reported by the APR2012 and shown in figure III.18 represent the total of families in each treatment group who benefitted: 953 in phase I and 372 in phase II. The report states that water systems were re-habilitated or constructed in all the bateys and that all the families have benefited from the water systems. The next figure compares the Treatment Groups on this indicator. The comparison has a high statistical significance, with a Chi-square (1) test yielding a value of 288.912, ($p = .000$). The analysis also yields a very large effect size ($w = 1.03$) and a perfect power (1).

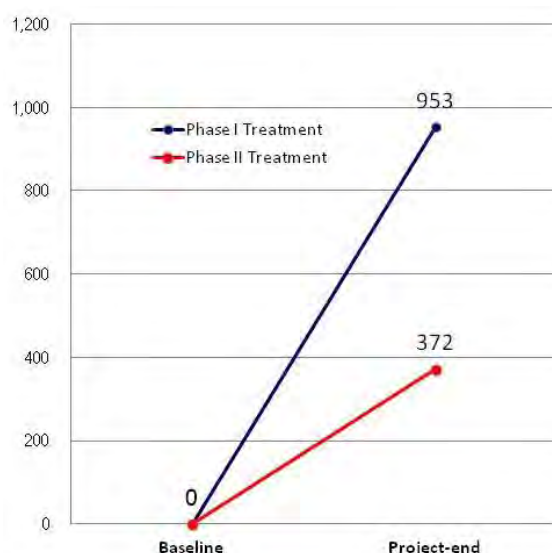


Figure III.18: # of families benefitting from improved water systems

2. Project intermediate outcomes

3.35 The present study has gathered evidence on how the delivery of the Project's development outputs has likely induced changes in the living conditions of the intervened batey communities. This evidence includes indicators that, while still measuring the delivery of outputs, can only reach noticeable values if some degree of cooperation or change in attitudes is present in the beneficiaries themselves. This is the case, for instance, of the delivered outputs: *women receiving PAP smear tests*, or *enrolled primary school students attending classes*. These are actions traditionally inhibited or hampered, in poor, uneducated communities, by a number of socio-cultural or economic constraints, such as taboos on women's intimacy or the pervasiveness of child labor. So, this type of outputs could not really be delivered in any significant way, if those barriers did not somehow begin to be overcome and some degree of concurrent proclivity or disposition on the part of the population is already at work. This, in the opinion of the present researchers, signals at least a trend toward attitudinal changes; and this is the reason why we deemed these outputs to fall right on the frontier with *intermediate outcomes* (early development effects of the Project). Intermediate outcomes observed in the intervened bateys of San Pedro de Macorís – Hato Mayor, as discussed below, signal changes in the living conditions of the population in areas of: (i) health [with improved maternal, child, and reproductive health; HIV/AIDS prevention; TB prevention and treatment; and hygiene among batey residents]; (ii) education [with better infrastructure, increased access to, and improved quality of, primary school education and

extracurricular educational services]; and (iii) sanitation [with improved water and asepsis systems]; as well with better housing and other services.

a. Health

3.36 The percentage of all of women between the ages of 15 and 49 in the bateys who have received at least one PAP smear test in the past 12 months, at the moment of measurement, is a first important indicator of intermediate health outcomes induced by the Project. Although the value reported in the APR2012 for the baseline of the phase II treatment group was 37%, in this analysis we use the value 39%, calculated directly from the previous studies' database. We confirmed that the statistical interpretation of the results was not affected by the use of this alternative value.

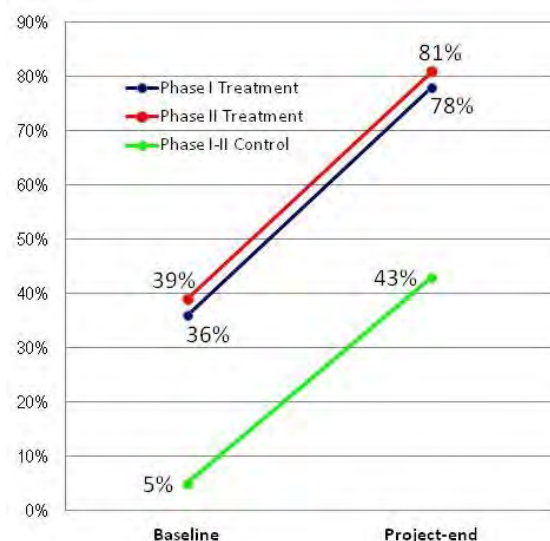


Figure III.19: % of women 15-49 receiving a PAP smear test in the past 12 months

3.37 Since this indicators is admittedly a late output indicator -accepted hear as “proxy” of an early outcome as well- the values reached in the treatment group are unmistakably due to the Project. However, these results cannot be clearly contrasted with results in the surveyed control group which, in terms of trend behavior, are almost exactly the same. (See figure III.19). The treatment groups registered very similar percentages of women receiving PAP tests at the baseline studies -36% for phase I and 39% for phase II- while the Project-end studies found that those percentages rose significantly to 78% for phase I and 81% for phase II. Now, the data for the control group points to almost exactly the same trend increase in the percentage of women receiving PAP test, despite the fact that the rise in the control group was from an very low baseline value (5%) and, as a result, the positive trend only caused such percentage to rise to 43% in that group: just above the level the treatment groups were already exhibiting before the start of the Project. The quantitative analysis shows that this discrepancy between treatment and control groups may not be attributed to mere chance, as the statistical significance of the difference is good ($\text{Chi-square}_2 = 9.199$, $p = .01$). Also the effect size is large ($w = .57$) and the **power** of the analysis is very high (1). However, the fact that trend results (i.e. the slopes of lines) of both treatment and controls group are the same sheds doubts on the conclusion that these results are really attributable to the Project.¹⁰

¹⁰ In this context it must be emphasized that, while statistical significance tests guard against the possibility of reaching conclusions due to mere happenstance or random errors, their reliability also assumes that underlying systematic measurement errors are not being made either. In discussing the above conspicuous results with the Project executors and supervisors, the possibility has been raised that one such systematic error might have been present concerning the reliability of the groups chosen for comparison at the start. Specifically, a cause for the identical slope between the changes in PAP smear test between control and treatment groups may have been that the two groups chosen (treatment and control) were in fact porous and not perfectly isolated from one another – arguably a non-random, systematic measurement error, but one nearly impossible to fully avoid in most social project cases. In such a situation, just a word of mouth could have spread the message of the benefits of receiving the test among women which, being of separate communities visited each other regularly anyway; thus inducing a positive “contagion effect” of sorts between communities in term of the PAP test practices, regardless of what project or institution administered the tests. However, no firm evidence supporting this circumstance is available.

3.38 Now, regardless of what is the real reason behind the intriguing results on the PAP smear test indicator, it is safe to state that we are here in the presence of a result that was not necessarily intended, but nevertheless positive even if not attributed to the Project; a kind of result stipulated as possible in our present evaluation framework -See *Note on Evaluability of the Project*, pages I and ii.

3.39 Another health intermediate outcome (also in the frontier with outputs) likely induced by the Project is reflected in the percentage of children under 5 years of age who received vitamin A from the Project. The information on this indicator is not included in the databases from the baseline study for treatment groups in both phases I and II, or in the APR2012. For this reason we cannot perform our main treatment-control comparison and we can only show the percentages found in the Project-end studies, as we do in figure III.20. In the SPM Control baseline we found 52% of children receiving Vitamin A, although the source providing the Vitamin is not specified. As can be observed, while percentages for the two treatment groups in the Project-end study reached close to 85%, in the control group the percentage of children under 5 years of age who received vitamin A from was only 68%. This difference does not, however, presents statistical significance ($\text{Chi-square}_2 = 2.454$, $p = .293$); so, statistically, it can be attributed to factors other than the Project's intervention.

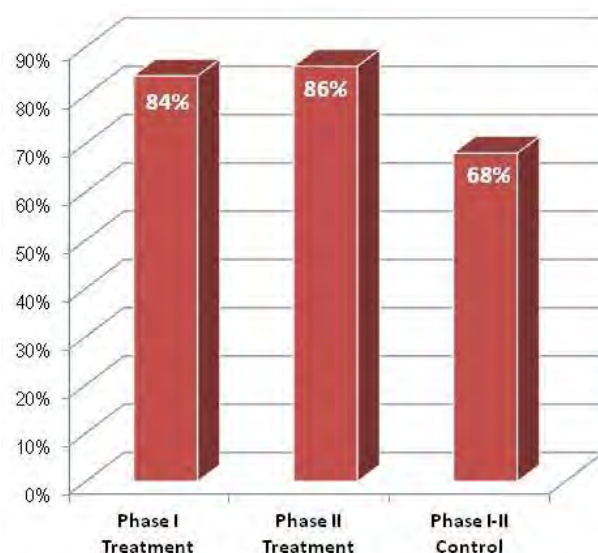


Figure III.20: % of children under 5 getting vitamin A.

b. Education

3.40 An important expected early outcome of the Project, in the area of education, was to be measured by the number of children enrolled in preschool. Yet, no evidence of a positive outcome induced in this respect could be found. On the contrary, a negative trend was detected, thereon, as shown in figure III.21; a startling result already discussed in the present report (see paragraph 2.27 above) although with no statistical significance. This indicator was only included in the Project-end study for the control group -where we registered 0 students enrolled in preschool- but there was no baseline data gathered for such indicator to compare it to in the control group. In the baseline phase I treatment group survey a question was included to estimate if every, some, or none of the children had received preschool education; but it is not possible to determine the actual number of children enrolled from this question.

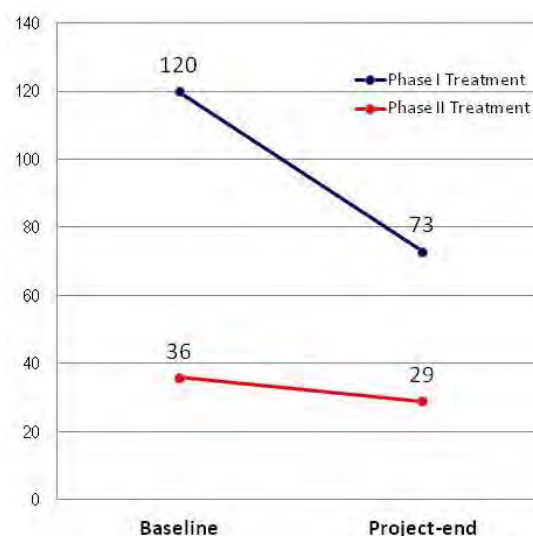


Figure III.21: % of children enrolled in pre-school.

3.41 According to the information on the APR2012, the number of children enrolled in preschool actually decreased by the end of the Project in both phases, and the differences did not reach statistical

significance either ($\text{Chi square}_1 = .938, p = .333$). So, once again, no likely attribution to the Project may be argued for the case of these results.

3.42 Another intermediate education outcome of the Project was to be measured by the number of youth and adults participating in vocational or informal education activities in the intervened batey communities. Again, this indicator was not included in the baseline control survey. In the Project-end control study we registered 0 youth and adults participating in vocational or informal education programs, but for the above reason we cannot compare this value with the one at the ex ante situation in the control group. The data we have examined on this indicator was supplied by the APR2012.

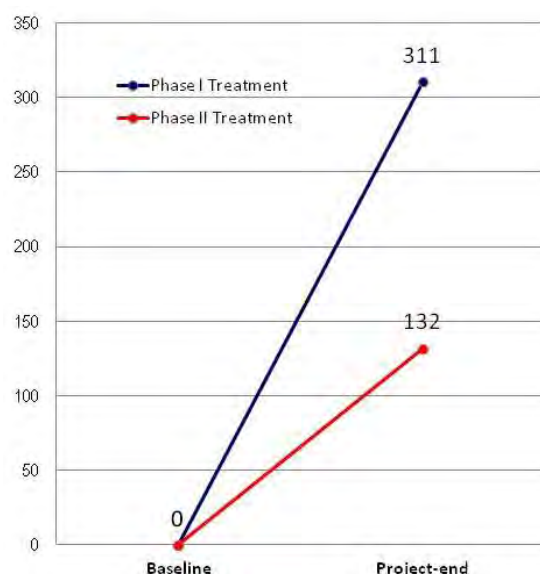


Figure III.22: # of youth, adults on vocational or informal education.

3.43 Now, as it illustrated in figure III.22, coming from a baseline of 0 participants in both treatment Groups of phases I and II, by the end of the project 311 youth and adults participated in USG supported vocational or informal education programs for phase I, and 132 in for phase II. Results of the Chi-square test associated with these differences showed a high statistical significance ($\text{Chi-square} (1) = 366.155, p = .000$). The effect size is very large ($w = 1.15$) and the power of the analysis also very high (1).

3.44 On the number of children and youth (6-14 years old) enrolled in extracurricular programs, the corresponding indicator also shows a positive early outcome likely induced by the Project. This indicator was not included either in the baseline control survey and in the Project-end Control study we registered 0 students enrolled.

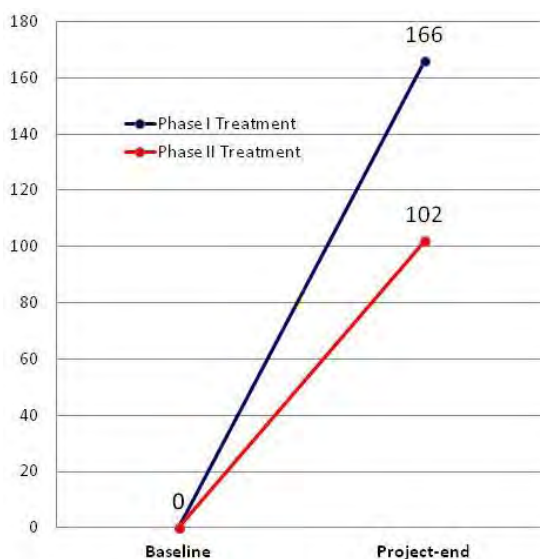


Figure III.23: # of children, youth on extracurricular programs.

3.45 From a zero enrollment in both treatment groups before the start of the Project, by the end there were 166 children and youth enrolled in USG supported extracurricular programs in the phase I bateys, and 102 in the phase II bateys. These programs included academic reinforcement and building of life skills. The data, which comes from the APR2012 report, is depicted in figure III.23.

3.46 Finally, in measuring the important early education outcome indicator: *percentage of enrolled primary school students who attend classes*, the actual comparison between treatment and control groups yields conspicuous results, also deserving special comments from researchers. The school attendance only augmented from 80% to 82% of students in the phase I treatment group, and in the phase II treatment Group it augmented from 74% to 85%, by the

end of the Project, as depicted in figure II.24. These represent modest, but clear improvements toward a better habit of class attendance by primary students. However, the primary school attendance in the control group appears to have risen not only from a higher mark of 97% but actually to reach the totality (100%) of enrolled students.

3.47 Researchers found that this curious result concerning the control group probably has a very probable explanation in the way the field data was gathered to measure the indicator. The data on this indicator, from the project-end studies on the treatment group hired by STC came from staff monitors reviewing attendance, each first Wednesday of the month in both phases I and II. This is a proper, objective way of collecting data on such indicator, but one that could only be applied when the schools were open. Since the present research's field surveys coincided with the school vacation time, no such method could be applied. Yet, since data for the baseline in the control groups was collected by the previous study on the basis of a question included then in the survey instrument, the researchers, for consistency purposes, decided to include the same question in the present survey as well.

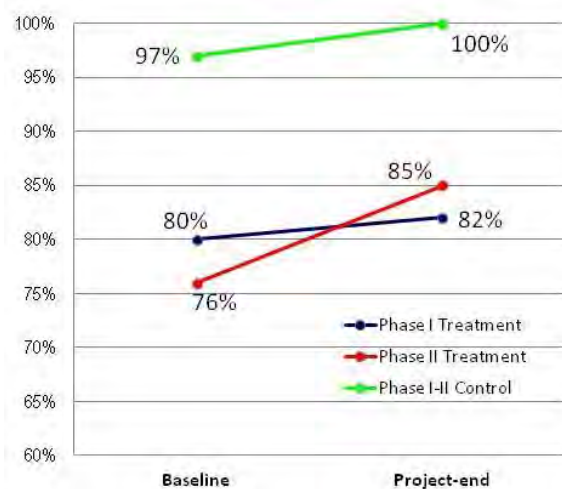


Figure III.24: % of enrolled primary students attending.

3.48 Accordingly, both the baseline and project-end surveys for the control group provide “interviewee estimations” for this indicator, and it is clear that -for reasons unknown to the present researches- such estimations are consistently higher than the numbers obtained by STC through more objective methods. So, because they are deemed less than objective, we consider those estimations of the interviewees on school attendance in the control groups to be not reliable. At any rate, the relevant differences turned out not to be statistically significant either; with the Chi-square₂ test yielding .337, and a *p* of .845

c. Sanitation

3.49 Perhaps the most important early change expected in living conditions relating to sanitation revolve around the access of batey families to clean drinking water. Here again, contrary to the situation on early outcomes in education discussed in the preceding section, changes of living conditions in sanitation could in part be corroborated during our evaluation surveys of families; especially in what concerns the actual drinking of clean water. Therefore, the researchers repeated in the said surveys a question on source of actual drinking water, already included in previous surveys, to validate the results obtained on the above mentioned indicator. These validation measures and the corresponding findings by the field research are discussed below, to complement the findings by the main indicator of access to clean drinking water.

3.50 The data provided by STC in the APR2012 measures the number of families with improved access to clean drinking water by the number of filters provided to the households (properly a project *output*). Bear in mind, though, that the *early outcome* implied here assumes that such filters are actually used by the families (a budding behavioral change); this being the reason for the validation questions discussed below. By the date of the APR2012 report, it indicated to have distributed 950 filters in the intervened communities of the phase I and 76 -out of a target of 300- in the intervened communities of phase II.

The comparison can be appreciated in figure III.25. The validation question about the actual *source for drinking water* included in the surveys yielded unexpected results, also deserving special comments. In order to compare the situation for the control groups, as with the previous indicator, we did an analysis of the answers to the said question.

3.51 As shown in figure III.26, in phase I there was a considerable increase of drinking water from bottles, from the baseline to the Project-end study; decreasing consequently the use from other sources, such as “Buy from a truck”. Only 5% mentioned to have a “Water filter installed in the household”. In phase II the increases from the baseline to the Project-end study were for “From bottle” and “A community-water tap”, decreasing other sources, such as “Well”, “Buy from a truck” and “A tap in the home yard”. Only 1% mentioned a water filter. In the Control studies, from the Baseline to the Project-end, there was also an increase for “From bottle” and for “A community-water tap”, and a decrease for “A tap in the home yard”. As it turned out, the water pumps that serviced the households were out of order in two of the control communities; so, it seems that this might have been the reason for the decrease in “A tap in the home yard” as a mentioned source of water. Now, the general results shown in the figure on the alternative “Water filter installed in the household” makes one wonder: what happened with the filters distributed by the Project?

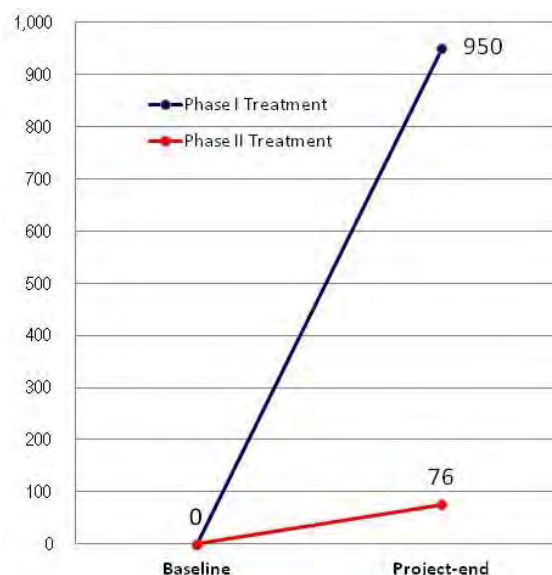


Figure III.25: # of families with improved access to clean drinking water.

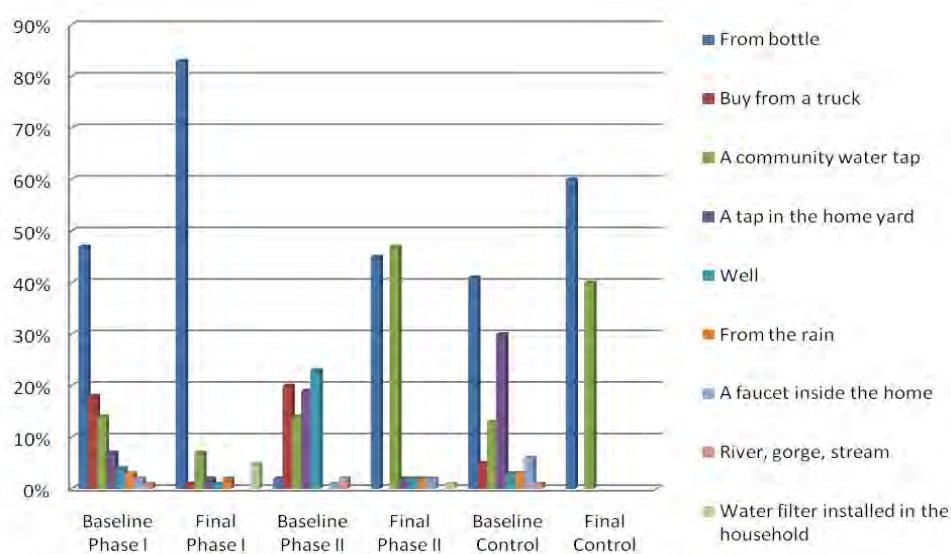


Figure III.26: Source of drinking water in the bateys

3.52 In the opinion of the present researchers, the apparent invisibility of the 1,026 filters distributed by the BCDP among the communities may be due to an inadequate formulation of the question. From the very first studies the actual question was: “From where do you regularly get you drinking water?” As it is not

possible to drink the water directly from the filters, people in these communities have to transfer the water from the filter to a 5 gallons plastic bottle. Then, when they want to drink some water, they take it from the bottle. In the houses with water filters installed, it is very probable that in answering the question many of the interviewees thought that the correct answer was “From bottle” and not “From the filter”. The question probably should have been a more direct one, like: Do you have a water filter installed in your household? Or: Do you use a water filter for drinking water in the family? However, consistency for comparison with previous studies precluded this precision in our present survey.

3. Project terminal outcomes

3.53 Although the BCDP evaluation premises and framework did not allow for a properly measurable examination of project impacts (long term, structural changes contributed to by the intervention), by the end of the Project some terminal outcomes were already visible in the bateys reality, in the form of incipient changes in patterns of conduct in the target population; the same which, if sustained, can be seen as precursor marks for longer term, more permanent transformations in the future. Being time-associated and aligned with the development outputs delivered and proximal development outcomes induced by the BCDP -and not explainable by pure chance- it can be argued that these nascent behavioral changes are likely attributed to the Project.

3.54 According to the structure of the intervention, changes in behaviors and associated living results have been detected mainly in areas of: (i) health and wellness practices; (ii) education payoffs and (iii) sanitation habits of the targeted batey population.

a. Health

3.55 One measure of budding new habits relevant in this area of development is whether parents or other adults responsible for the well being of children can report to be applying key child health and nutrition practices learned through the Project. On this respect the Project has clearly made a difference, as depicted in figure III.27. In every study, interviewees were asked about practices intended to promote their children’s health and nutrition. Then, the answers were analyzed and classified between adequate and inadequate, based on the pre-arranged set of recommendations provided by the Project. The adequate answers were then added in the databases, for the calculation of the individuals able to actually identify and apply adequate practices to protect the wellness of children in the bateys intervened by the Project.

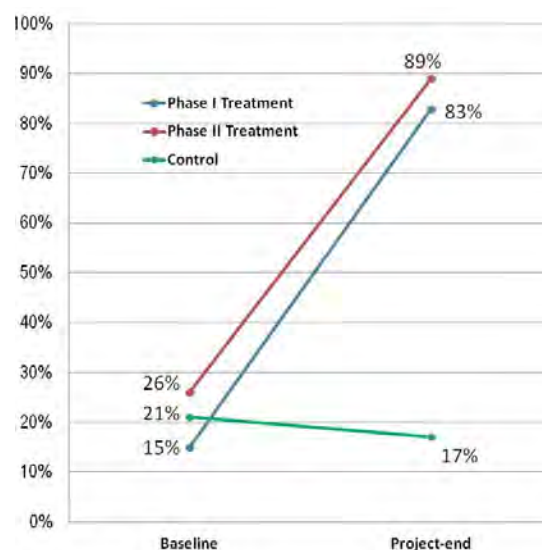


Figure III.27: % of care givers reporting 3 good child health, nutrition practices.

3.56 As we can see by the *percentages of care givers who report at least three key positive child health and nutrition practices* shown in the figure, the Project appears to have had a distinctive effect on the child health and nutrition practices in the Treatment communities. In the Baseline studies the percentage of care givers reporting three or more positive practices was very similar in the three groups studied, ranging from 15% to 26%. By the end of the program, the percentages of these caregivers climbed to

83% in the phase I Treatment Group and to 89% in the phase II Treatment Group, while they kept relatively flat in the Control Group, at 17%.

3.57 Another wellness promoting behavior expected as advanced outcome of the Project is the practice of breast feeding as an exclusive means of baby nurturing. The information on the indicator: *% infants under 6 months of age exclusively breastfed within the past 24 hours* at the time of measuring was not included in the database of the phase I treatment group, or in the APR2012. This limits us in the application of our main treatment-control comparison only to the data on the Project's second phase.

3.58 As is illustrated on figure III.28, departing from a baseline of 22% in the phase II treatment group, the percentage of infants under 6 months of age exclusively breastfed within the past 24 hours climbs to 50% by the end of the Project. The phase I treatment group registered at the end 53%, from an unknown baseline. This percentage stays without visible changes in the control group, which suggest a sound effect of the Project in the treatment group. Calculated upon the Project-end studies only, because of the lack of information for the phase I treatment group, the statistical significance of these differences turned out to be high, with a Chi-square₂ yielding a value of 51.631 and a *p* of .000. Also in this case the effect size is large (*w* = .71) and the **power** of analysis is very high (.99), confirming that the differences found between treatment and control groups can be attributed to the Project's intervention, at least for the bateys of the second phase.

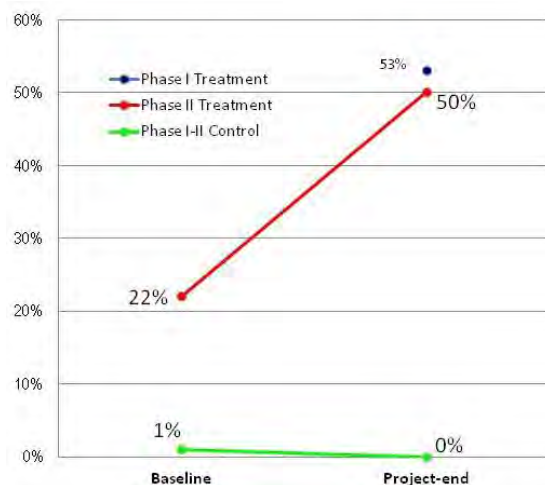


Figure III.28: % of infants under 6 months exclusively breastfed in the last 24 hours.

3.59 The Project also promoted practices of reproductive health, especially amongst females in the intervened bateys. The development outcome in this respect was to be measured by the *percentage of women ages 15-49 reporting two key reproductive health practices*. As with the indicator about child health and nutrition practices, women in this age bracket were asked for their key reproductive health practices. Their answers were analyzed for their adequateness and the proper responses were added in the databases. The percentages of women between the ages of 15 and 49 who could identify and report at least two key reproductive health practices are shown in figure III.29.

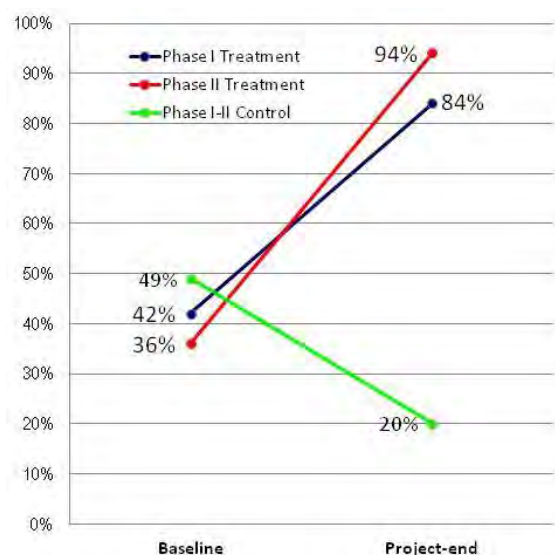


Figure III.29: % of women 15-49 who report two key reproductive health practices.

3.60 Again, as we can observe in the figure, the Project had a distinctive effect on the reproductive health practices of women in the intervened communities. In the baseline studies the percentage of women in reproductive reporting with two or more of the said positive practices was very similar in the three groups studied, ranging from 36% to 49%. The Project-end studies showed that the percentages of these women rose to 84% in the phase I treatment group

and to 94% in the phase II treatment group, while it descended to 20% in the control group. These differences are statistically significant (Chi-square (2) = 38.387, $p = .000$), the effect size is medium ($w = .37$) and the **power** of analysis is very high (.97).

3.61 The prevention of STD and HIV/AIDS is another area in which the Project has induced statistically significant development outcomes. As in the indicator on women's reproductive health practices, the interviewees were asked to report their STD and HIV/AIDS prevention practices and the sum of the positive practices were added in the databases. Figure III.30 presents the result of this calculation in terms of the percentages of heads of households who identify at least two key STD and HIV/AIDS prevention practices for all the groups. The very sound effect of the Project in terms of this indicator is evident from the fact that starting from a common departure at the baseline, in the range of 34% to 38%, the percentage of household heads who identified at least two key STD and HIV/AIDS prevention practices grew to 87% for the phase I treatment group and to 99% for the phase II treatment group. The fact is also very salient that having started slightly below those of the treatment group (34%), the corresponding values in the control group fell markedly to a 14% in our Project-end survey. The statistical significance of the differences is high (Chi-square₂ = 37.829, $p = .000$); the effect size is medium ($w = .38$) and the **power** of the analysis is very high (.95). Therefore, the observed gap in percentages and trends between treatment and control groups concerning this indicator cannot be attributed to chance and is, arguably, imputable to the workings of the Project.

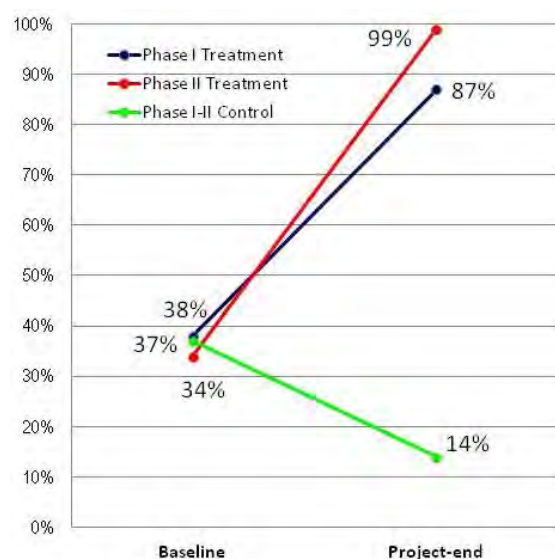


Figure III.30: % household heads identifying two STD-HIV/AIDS prevention practices.

3.62 Arguably more important than the health practices in adults, are the wellness promoting habits that the Project was supposed to induce in batey children; as these habits will presumably be carried with them through adulthood and transmitted down to incoming generations. The outcome consisting of school children exhibiting improved hygiene practices was to be measured by the indicators: (i) % of school-children who wash their hands after using the school latrine/bathroom; and (ii) % of school-children who wash their hands before eating the school breakfast/snack/lunch. Unfortunately, neither indicator was included in the databases from the baseline studies. Therefore, we are limited to present the results for this indicator only from the Project-end studies, as in figure III.31. This report is based on direct observation at the schools.

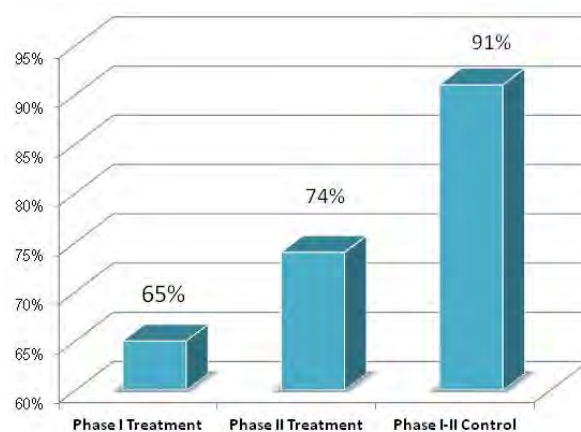


Figure III.31: % of children washing their hands after using school latrine/bathroom.

3.63 According to the data collected, the percentage of children who wash their hands after using the school latrine/bathroom appears doubtfully high in the control group (91%) at the end of the Project, as compared with a 74% in the phase II treatment group and a 65% in the phase I treatment group. This result may be related to the fact that the values reported in the APR2012 come from direct observation in the schools while the control group value comes from an estimation of the interviewees. Here again, as in the case of the indicator of primary school attendance, this is a proper, objective way of collecting data on the behavior of children, but one that can only be applied when the schools are open. This was not the case when the present research's field surveys took place; and so the data was collected directly from the answers to the questions posed to those present in the households. As affirmative answers to this question carry a high social desirability, the respondents could have estimated an inflated number of children washing their hands in this situation. Anyway, the differences among the groups were not found to be statistically significant: $\text{Chi-square}_2 = 2.869$, $p = .238$.

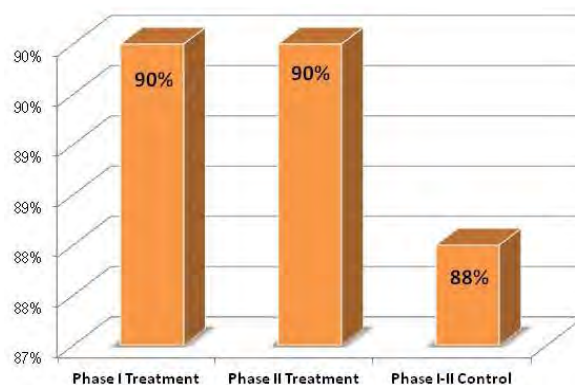


Figure III.32: % of children washing their hands before eating at school

3.64 Results found on the other component of the indicator (percentage of school-children who wash their hands before eating the school breakfast/snack/lunch) ended up with the same lack of significant differences. The only value reported from the APR2012 at the baseline was for the phase II treatment group, at 74% of students following the practice; but for the two treatment groups at the end of the program it was reported a 90%. In the control group, based on the interviewees' answers 88% of the children washed their hands before eating at school, as depicted on figure III.32. The same observation as above, on the different source of the data -direct observation at the school vs. interviewees' estimation- applies in this case; and so, results from the control group survey are considered not reliable in the present context. As mentioned, the differences among the groups lacked statistical significance anyway ($\text{Chi-square}_2 = .03$; $p = .985$).

3.65 Perhaps the development outcome where the Project went farthest afield in the area of health is the one concerning the incidence of diarrhea in small children: an outcome probably already in the frontier with the realm of impacts (i.e. arguably, an *advanced effect* starting the path toward longer term *impacts* such as lower infant morbidity and/or mortality in the batey communities).

The stipulated metrics for this outcome is the % of children under-5 with diarrhea in the last two weeks at the moment of measurement, and the values found are depicted in figure III.33. Most of the values reported were calculated from the different databases of previous baselines and project-end studies, unless otherwise indicated. The value of the Project-end control group was, as usual, calculated based on the data collected in our own evaluative survey.

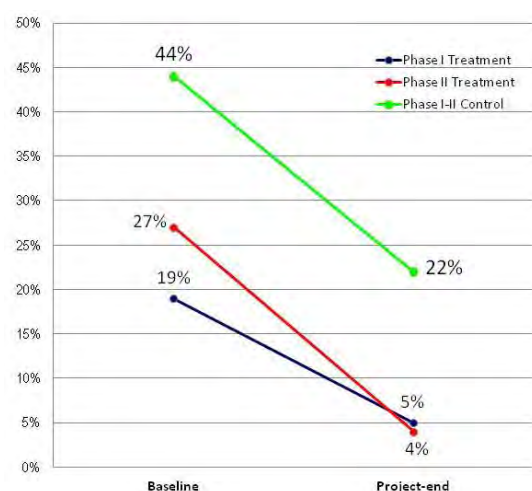


Figure III.33 % children under 5 with diarrhea in the last two weeks

3.66 As can be seen, the incidence of diarrhea in children under five at the baseline was 19% in the phase I treatment bateys and 27% in those of phase II. These percentages fell to 5% and 4% by the end of the Project, revealing a noticeable change associated the intervention. However, the attribution of this result to the Project can be doubted, because the same falling trend is observed in the Control Group. Here again, the incidence of diarrhea was so high at the baseline of that group (44% of children) that after the fall it only descended to 22%, a value within the baseline range of the treatment groups for both phases.¹¹ In any case, the differences found are statistically significant with Chi-square₂ value of 8.374 and a **p** of .015. The effect size is medium (**w** = .33) and the **power** is high (.85).

b. Education

3.67 The indicator stipulated to measure Project's terminal outcomes in the area of education -*% of students enrolled in 4th grade that pass 4th grade*- does also arguably borders the domain of impact measures; i.e. it gauges a terminal outcomes which can also be viewed as an early impact, presumably starting the path toward higher rates of graduation up the education ladder and longer term impacts such as, may be, employability in the future for batey residents. Here again, it was inappropriate to include questions on this subject in the surveys for the baseline and project-end control groups, and we analyzed only the information gathered by STC directly from the schools as presented figure III.34. The percentages of students enrolled in 4th grade who pass 4th grade went from a baseline of 84% in the phase I treatment group, and 88% in the phase II treatment group, to 95% in both groups by the end of the Project. Another terminal outcome worth measuring in the area of education, which is also bordering the early impacts domain, is extent to which grade 1 to 4 students were capable of reading at their grade level. Sadly, it appears that those responsible to do measurements, in order to calculate this indicator did not do the task. Thus, STC did not supplied information on this indicator in the APR2012.

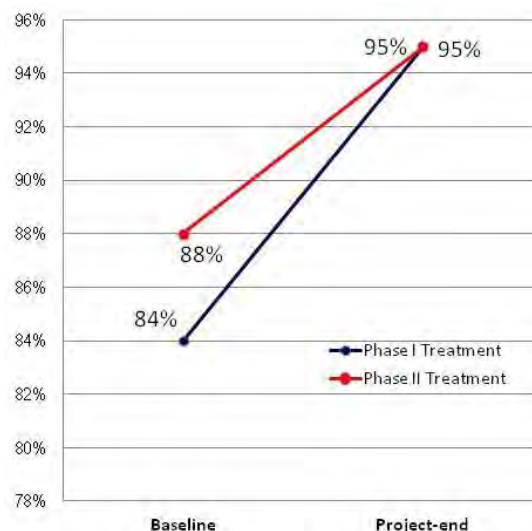


Figure III.34: % of students enrolled in 4th grade passing 4th grade.

a. Sanitation

3.68 A first indicator we use to validate the actual use of the installed water systems in the community was designed to obtain information on the sources of water actually used in the homes. So, in order to make possible a comparison with the control group, we asked the heads of households interviewed in this group the same question on the source of water for domestic use in the household, which was included in the studies for the treatment group in both phases of the Project, and both at the baseline and project-end points. Based on the frequency of answers, the percentage was calculated per category of a range of water sources previously defined.

¹¹ Similarly to what was discussed on par. 3.37 about the *PAP smear test indicator*, the two explanations deemed more likely for the present results are: either that they are not attributable to the Project at all; or that there is a positive "contagion effect" in the control group, due to its permeable nature with respect to the treatment group.

3.69 Figure III.35 summarizes and compares the resulting percentages and their distribution between the eight identified use categories, for every group in each phase, and both at the baseline point and at the project-end point.

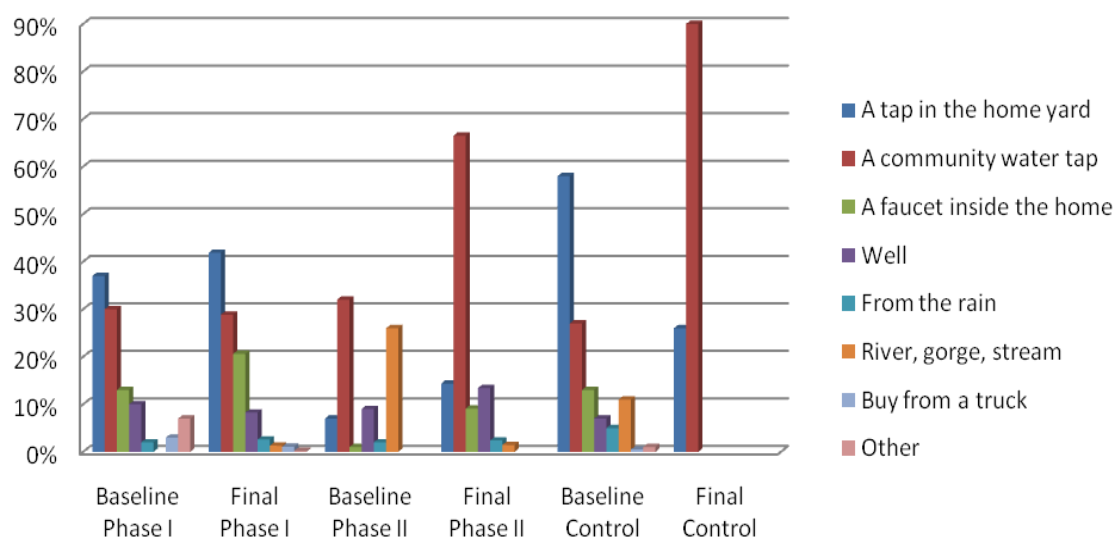


Figure III.35: Source of water for domestic use

3.70 As can be observed in the figure, for the phase I we cannot see a major difference in the values of the indicators between the baseline and the project-end studies. The most used water source was “A tap in the home yard” in both cases. The second place was for “A community water tap”, and the third place for “A faucet inside the home”. During the quantitative analysis of this particular set of results, the significance of the differences found could not be statistically established because of the many empty cells in the comparison. In phase II there is an increase from the baseline to the end of the in the percentage of families using a community water tap and a decrease in the percentage of families using water from a river, gorge or stream, showing an improvement associated with the water systems installed by the Project in the treatment group. However, the same increased use of the community tap as source of water for domestic use is evident in the families of the control group. Indeed, this group’s results show a majority using “A tap in the home yard” in the Baseline, and in the Project-end study 90% of the interviewees mentioned “A community water tap”. This vast majority is most probably due to the fact, reported by our field supervisor, that the water pumps servicing the houses were out of order in two communities at the time of the survey.

4. Other project results

3.71 Additional Project results in the batey communities of SPM-Hato Mayor refer to the areas of housing, emergency preparedness and income generation.

a. Housing

3.72 The stipulated indicator to measures this Project output was the *# of families with improved housing*. The Treatment values reported in the APR2012 indicate that the number of houses rehabilitated in a major way or constructed by the Project in the treatment groups was 600 for phase I and 176 for the phase II. In the area of housing the Project not only did major rehabilitation or full construction, but also

repairs. The number of all family houses that were improved upon and the percentage of all houses that were repaired by the Project in the intervened bateys is shown in figure III.36 and III.7, respectively.

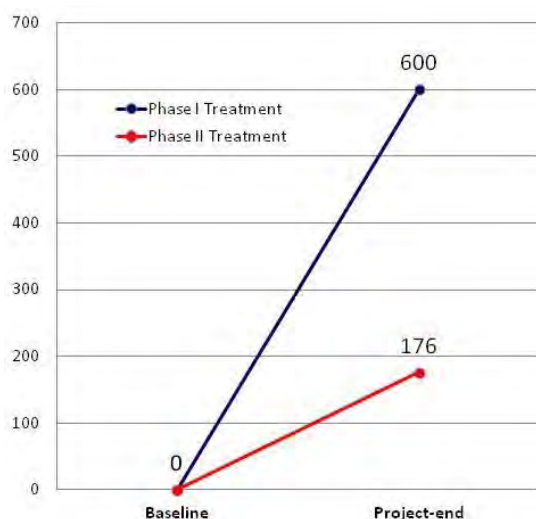


Figure III.36: # of families with improved housing

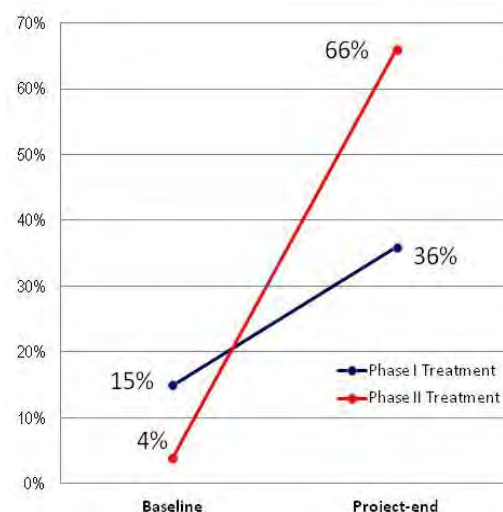


Figure III.37: % of houses repaired

3.73 To validate the information supplied by STC we analyzed a question included in the previous surveys to estimate the proportion of houses that have been repaired, based on the interviewees' perception. Unfortunately, this question was not included in the baseline study for the Control Group and, so, we cannot make comparisons with results from the project-end control survey, which was 0 houses repaired. Now, based on the perceptions of the interviewees, the comparisons within the treatment groups tend to validate the values of the house improvement indicator. As we can see, the percentages of repaired houses were very low in the Baseline studies: 15% for Phase I and 4% for Phase II. In the Project-end studies the repairs increased to 36% in the Phase I and to 66% in the Phase II.

b. Emergency preparedness

3.74 The area of preparation of communities to face emergencies included both a physical output and organizational activities, under the Project Enabling Strategy based on participatory community mobilization.

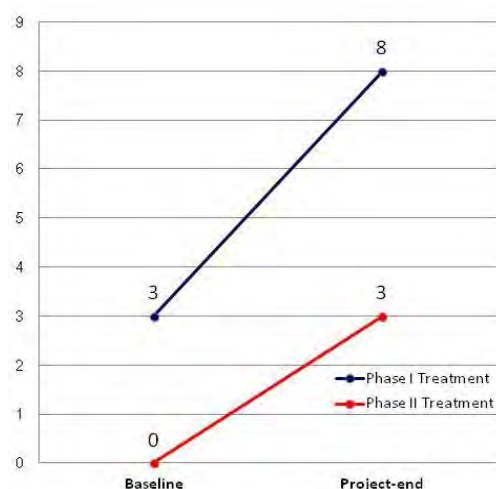


Figure III.38: # of bateys with access to an emergency safe structure

The physical output refers to the provision by the Project of safe structures (community centers) for use by the community as a refuge in case of an emergency. The organizational side of the contribution involved the development and launching of emergency preparedness plans for use by both the communities at large and the schools in particular. Figure III.38 depicts the output results delivered by the Project in terms of # of bateys with access to an emergency safe structure. The information found in the APR2012 indicates that, by the end of the Project, all the eight communities in the phase I treatment group were provided an emergency safe structure. Of the 4 communities in the phase II treatment group, only 3 did receive such a structure.

3.75 Researchers did not find any emergency safe structure in the Project-end control survey and there is not available information on the situation at the moment of the Baseline Control study. The comparison among the Treatment Groups based on the different studies can be appreciated in the figure. Because of the reduced samples, the statistical analysis of this comparison is inadequate, but its appearance coincides with the general direction of an appropriate delivery by the Project.

3.76 Concerning the existence of updated emergency plans for the community, the # of bateys with such plans is shown in figure III.39. According to the APR2012 all the bateys in the Phases I and II Treatment Groups had an updated emergency plan by the end of the project. In our quantitative Control evaluation survey we found that every interviewee considered that their community did not have an emergency plan, but this information was not included in the Baseline Control survey and that is why we limit this comparison to the Treatment Groups in the figure. For the same reason as in the previous indicator, the statistical analysis is inadequate; but again, the direction of the results coincides with our general hypothesis of appropriate delivery by the Project. In what concerns the Project output in the area of participatory emergency plans for schools, the results are shown in figure III.39.

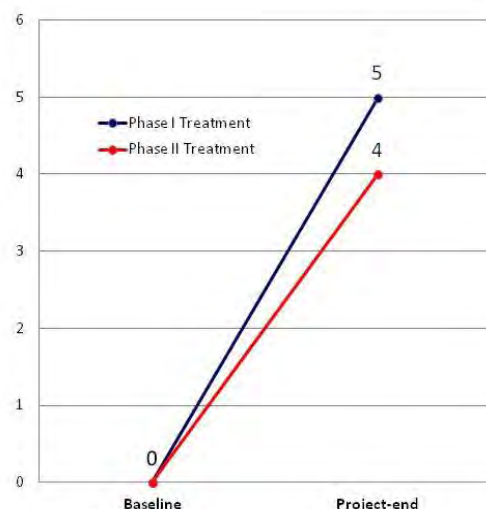


Figure III.39: # of bateys w/updated emergency plans.

3.77 In what concerns the batey schools having self-managed plans for collective protective actions in case of emergency, according to the APR2012 report, 5 schools in the Phase I Treatment Group and 4 in the Phase II had such emergency plan and were furnished with appropriate response materials and equipment, in case of implementation of such plan. There is no information available on the base line for this indicator from the control groups and, as with other indicators previously discussed in the present report, our comparison has to be limited to results observed in the treatment groups of both phases of the Project. Also, as in the previous indicator, the small sample impedes to perform a statistical analysis but the direction of the results goes even with our hypothesis of appropriate delivery by the Project. Results on this indicator are presented in figure III.40.

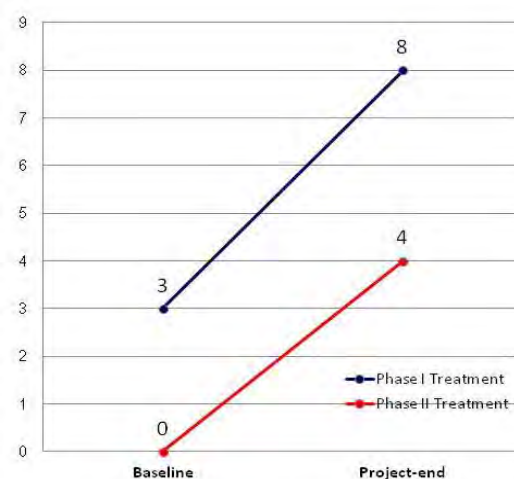


Figure III.40: # of batey schools with updated emergency plans.

c. Income generation

3.78 The stipulated indicator to measure Project outputs in this particular area of development intervention is the # of batey residents who have actually received some support in income generation activities. The APR2012 reports that 222 people, from the treatment communities in phase I have, by the end of the Project, received support in income generation activities such as the training by the Project of individuals in *home gardening*. This activity was intended in part for the purpose of family consumption,

thus complementing nutrition inside the households. House gardening activities were also intended for selling crops to outside customers and, thus, complementing the family's income. Finally, the Project's activities on this respect also included the provision of loans to finance micro-business activities in the intervened batey communities, as a way to complement the incomes of participating families. In Phase II, 83 people received this help in the Treatment communities. The results obtained by the study are depicted in figure III.41

B. Quantitative perspective 2: Other bateys and regions of phase II

3.79 Given the severe limitations in terms of the number of indicators actually measured in the four bateys of the regions newly included since 2011 in the BCDP intervention, as compared to the amount data collected on Project results in the region of San Pedro de Macorís – Hato Mayor, there is very little evidence of Project results that can be discussed; and the scant numbers available refer only to Project outputs. No determination could, therefore, be made in either of the two new regions (Verón-Bávaro and Bahoruco) about possible intermediate outcomes, terminal outcomes or impacts of the Project. Based on the data, there is also no need and use for statistical significance tests on results in these bateys.

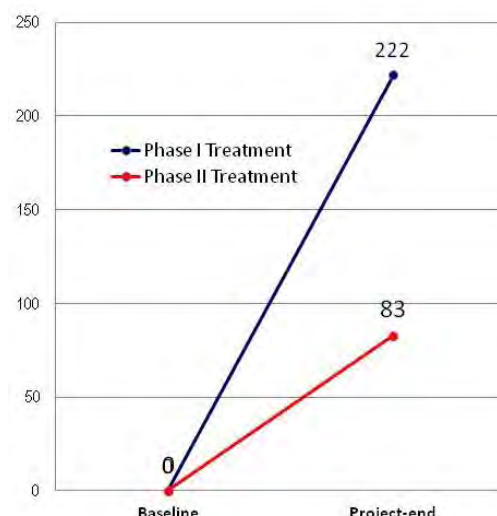


Figure III.41: # of residents receiving support in income generation activities.

1. Results in the region of Verón-Bávaro

3.80 The field data available for the indicators applied by the Project in the bateys of the Verón-Bávaro region is summarized in the table III.1.

Indicator	TREATMENT GROUP		CONTROL GROUP	
	Baseline	Project-end	Baseline	Project-end
Number of people benefitting from USG supported health services	0	9,604	0	0
Number of Community Health Promoters trained and equipped	0	8	0	0
Number of families with access to improved sanitation services	0	208	0	0
Number of families with improved access to clean drinking water	0	10	n/a	n/a

Table III.1: Comparison of field data on results of the BCDP in Verón-Bávaro .

a. Number of people benefitting from USG supported health services

3.81 The APR2012 reports that the number of people benefitting from USG supported health services in the VB zone is 9,604, due in part to the establishment of a close relationship with the Ministry of Health (MOH). Coming from no health services, now the communities have weekly visits from doctors and

nurses and the MISPAS is building a permanent clinic in *Hoyo de Friusa*. Our evaluation survey in VB confirmed that there were no beneficiaries of the project among the interviewees of the control group.

b. Number of Community Health Promoters trained and equipped

3.82 As in other regions, the BCDP trained and equipped community health promoters in Verón-Bávaro on the subjects of: (i) Maternal/newborn health; (ii) Child health and nutrition; (iii) Reproductive health/family planning; and (iv) TB management. As in the SPM region, we analyzed these indicators together as the values are the same for both. The Treatment information was supplied by the APR2012, and our field survey confirms no community health promoters operating in the control bateys.

c. Number of families with access to improved sanitation services

3.83 In the corresponding section referring to the SPM–Hato Mayor micro-region we analyzed additional data from the different surveys to reflect the perception of the interviewees on the sanitation services in the communities and to validate the data supplied by STC. However, in the micro-region of Verón-Bávaro we found no data from the treatment group at the end of the Project, because there was not a project-end study in this geographical area that included the desired validation questions.

d. # of families with improved access to clean drinking water

3.84 As mentioned in the SPM – Hato Mayor section, the values of the APR2012 indicate the number of water filters installed in households of the communities. Because of operational delays, the Project could only distribute 10 filters in the Verón-Bávaro zone. As we don't have information on this indicator for the control groups at the baseline, the comparison is limited to results on the treatment group.

e. Other results

3.85 No end-project treatment study was performed in the region on any other indicator included in the *Final List*, so no evidence of additional field results of the Project can be discussed

2. Results in the region of Bahoruco

3.86 The field data available for the indicators applied by the Project in the bateys of the Bahoruco region is summarized in the table III.2.

Indicator/area	TREATMENT GROUP		CONTROL GROUP	
	Baseline	Project-end	Baseline	Project-end
Number of people benefitting from USG supported health services	0	1,845	n/a	n/a
Number of Community Health Promoters trained and equipped	0	4	n/a	n/a
Number of families with access to improved sanitation services	0	50	n/a	n/a
Number of families with improved access to clean drinking water	0	12	n/a	n/a

Table III.2: Comparison of field data on results of the BCDP in Bahoruco .

- 3.87 Comparisons of all results data in this region have to be limited to the information provided by the APR2012, including the reference to the 1,845 people benefitting from USG supported health services, because there were no *Baseline* or *Project-end* study for the treatment group, or a *Baseline* study for the control group either. Furthermore, as we indicate in the section of validation of the beneficiary communities in the statistical appendix attached to the present report, and based on our evaluation survey, the Bahoruco communities of *Mena Abajo* and *Los Blocks de Mena*, could not be considered as control bateys because it turned out that they were actual beneficiaries of the Project, even though in the Project documentation they were classified as belonging to the control group,.
- 3.88 As before, the two indicators referring to Community Health Promoters are analyzed together because the values are the same for both of them. Although the direction of the results coincides with our hypothesis on the positive effect of the project, the small samples make a statistical analysis inadequate. Concerning the # of families with access to improved sanitation services, the value reported in the APR2012 indicates that 50 latrines were built in the intervened bateys of the region, but we do not have information to compare it with any other group. With respect to the number of families with improved access to clean drinking water, the value reported in the APR2012 indicates that 12 water filters were installed in households of the communities. Yet, as in the previous indicator, we do not have information from any other group to compare it with.

C. Overall qualitative perspective

- 3.89 The study included in-depth interviews with project managers and executors on more qualitative aspects of the Project, as well as with community leaders, members and field workers to validate issues of development approach, social methodology and quality of results, especially in what concerns the practical implementation of the Project's cross-sector enabling strategy, integrated approach to community development in the field, and the development of partnerships with other public and private actors. Specialized opinions were gathered from meetings with members of the management teams of STC and Mude and corroborating points as well as direct impressions were obtained from batey residents in intervened communities of San Pedro de Macoris, Hato Mayor and Verón-Bávaro.

1. Integrated approach to community development

- 3.90 Exchanges in the field with community leaders, ordinary residents and field worker corroborate that in all interventions the Project followed an approach to community development based on: (i) eliciting collective self-management and social mobilization; (ii) promotion of community consensus on needs and necessary actions; and (iii) practical study and application of ways to induce behavioral changes in individual and groups. This field approach, articulated by Project managers during the interviews, appears to have been instrumental in strengthening grassroots organizations, social cohesion, awareness of development issues -especially among batey women- and community initiative in solving identified needs. A four-stage process has been detected in the application of this Project approach to community development, including:
- (i) Initial focus groups to conduct self-studies on community problems and needs;
 - (ii) Formation of participatory groups with mutually shared interests to receive training and pursue sustainable solutions to the problems identified;

- (iii) Consolidation of the participatory groups; and
- (iv) Follow-up on actions.

3.91 The approach emphasized grassroots participation at all points of the process as well the combination of specialized training on specific subjects such as disease prevention, initial and vocational education, etc. with the actual delivery of goods, services and tools for the benefit of residents in areas such as infrastructure for health, education and sanitation, with the explicit goal of becoming self-managed by the community for future sustainability. This combination, also stressed by project managers in the interviewees, appears to have been successful in conveying the importance of social organization and mobilization to guarantee maintenance and sustain results, and for the general effectiveness of the Project. The quantitative data has provided enough empirical evidence of results associated with this emphasis on self-managing maintenance made by the Project. Prominent among the numerous cases documented in past chapters of clear deterioration in the not intervened bateys (control), as compared with the progress in those which were intervened, is the example of the good operation of latrines increasing through the year in the latter bateys, while in the control communities such operation has been deteriorating through time: a clear difference made between the two groups in terms of sustainability of results, by the concept of maintenance promoted by the Project. However, concerning possible synergies with external actors to foster sustainability of results from this integrated approach to community development, no evidence could be gathered, or opinions collected from interviewees, that the Project was particularly successful in creating partnerships with key external institutional actors, beyond some coordination with public entities necessary for the delivery of certain products.

2. Community organization & empowerment

3.92 With exceptions, in most intervened bateys researchers detected the presence of community organizations which were functional and active, and were launched mainly due to the Project induction and promotion activities. These organizations were: water committees; health committees; education committees; emergency committees; and women organizations; as well as organizations for the maintenance of the Community Center; youth clubs and energy committees. Members showed they know with precision what the organizations do and strive for in each case, and speak fluently about their frequent meetings and activities.

3.93 Residents and grassroots leaders confirm the influence exerted by the Project in motivating and/or reviving social organizations in the different treatment communities. Often they bear witness to the fact that organizations which existed before the start of the operation, but were lethargic or not functional at all, such as residents associations; agricultural associations; or parents-teachers associations were revitalized after the Project started and the importance of integration and self-management began to be conveyed in the bateys. Project influence in empowering groups and individuals at the grassroots level is particularly visible in the workings of disaster mitigation and response committees, local water boards and women's groups currently operating in the bateys under study.

3. Grassroots participation & self-management

3.94 Field interviews reveal that the intervened communities remain fairly involved in developing viable solutions to their self-identified problems and needs. Collective action as the one displayed for community works during the Project in construction of the community centers, and house repair and

refurbishing, is still frequent, in areas such as cleaning and waste disposal, water system repair and maintenance. Residents and leaders report that this type of activities are more frequent now than in years past, before the Project, when only political mobilization in times of elections were mostly the norm; or repairs and re-constructions quickly deteriorated for lack of community attention and maintenance.

- 3.95 An area of frequent mobilization often mentioned in the field in-depth interviews is the all important maintenance function of the local water boards and the contribution of community labor to the revamping of water systems. Yet, community actions in the pursuit of other goals are not infrequent, including collective petitions to authorities. A significant plurality of interviewed residents rates as *high* or *very high* the level of participation of people in the affairs of the community, although some also rate it as *low*. Also, a majority of individuals interviewed rate the success of the community mobilization as *High*, though some in the minority rate such success as *low*.
- 3.96 Even in some cases of communities with an appreciable level of mobilization, leaders report that the Project prepare them better for collective action; both aimed at promoting the welfare of individuals and families (such as through training on health preventive practices) and the welfare of the community as a whole. A particularly appreciated qualitative outcome of the BCDP in the intervened community is the universal agreement among interviewed community residents and leaders about the positive influence of the Project in lessening internal community violence and conflict. Some went as far as expressing that the BDCP all but eliminated the water conflict in their batey communities.

4. Cross-sectoral enabling strategy

- 3.97 There are numerous palpable benefits from the several interventions of the Project that are recognized by the interviewed batey residents. Prominent among them are the ones relating to the area of Health, where the role of the children *vaccination card* and the work of the community health promoters in charge of the *community case management* and in promoting PAP tests and family planning, as well as preventing of VIH, TB, Dengue and Cholera, has been highly valued by people. However, benefits from the enabling strategy implemented by the Project in the intervened communities are also recognized and appreciated as discussed below. Also, concerning the development of partnerships between the Project and public and private actors involved, as part of the enabling strategy for sustainability, the residents perceive and appreciate the useful coordination between the actors and its benefits to the communities. In what concerns the private sector, a particularly successful case of cooperation was the coalition USAID-Rotary International H2O Alliance to bring access to water and improve the sanitation, hygiene and health conditions of batey residents in phase II of the Project. A quarter of the additional resources used in that phase came from that partnership. However, a long term and close cooperation between the Project and possible partners in the public sector did not rank very high in people's perception; especially in judging the dedication of possible official partners in promoting development and welfare in the communities. For instance, interviewed residents were able to discriminate between the level of dedication and quality of work of the Project's health promoters and those of the health promoters that were recognized or appointed officially by the Ministry of Health; which hints at a limited synergy achieved with the public sector. Finally, the appreciation of residents is particularly noteworthy for the emergency plans enacted by the Project, together with the provision of safe infrastructure and equipment for in preparing for emergencies and responding to disasters. People spoke knowingly during interviews about the nature of such emergency plans and of specific cases in

which they have been practically launched during the last year or even within the last month at the time of the interview. Some mentioned, however, that the refugees are not appropriate or are already damaged.

- 3.98 An area raking lower in the general public appreciation of Project benefits is the area of *income generation* activities. Such activities are recognized as present and beneficial to people, both in what concerns family garden plots to complement household nutrition and incomes, as well as in relation to small-loans granted for micro-businesses in the intervened bateys. Yet, many declare not knowing or being benefitted from such activities. Finally, interviewees generally recognize the important coordination achieved with public institutions to organize the delivery of certain project outputs, such the launching of MISPAS health reference systems in the intervened communities, and the training of batey school teachers, based on contents approved by the authorities of education –as in the case of training in Literacy- this, despite the almost inexistent positive responses from residents and grass-roots participants in the community organizations, and project activities, about the meaningful development permanent partnerships with public institutions involved in the bateys for the promotion of sustainable development results.

IV. Analysis and conclusions

- 4.1 In what follows we mainly discuss what payoffs has been obtained by the BCDP development efforts in the intervened batey communities, based on the analysis of the present study's findings. Specifically, we summarize the answers obtained to our initial research questions, referred to: (i) the effect of the Project on the quality of life in the batey residents; (ii) the repercussions of Project on the participation, organization and mobilization of the batey communities; and (iii) the extent to which the Project has been successful in creating partnerships with key external actors. Since the evaluation metrics defined at the start of the operation were restricted to answering the first of the above questions, the resulting measurement and quantitative testing of the research hypothesis focused only on that question. In turn, with no pre-defined metrics and data available that could be used in the study's adopted quasi-experimental design to address the second and third research questions, those were tackled through a more qualitative appraisal of results, also included in the discussion below.

A. Effect of the Project on the batey residents' quality of life

- 4.2 According to our evaluation premises, the BCDP underlying *project theory* (development hypothesis) was that the batey's quality of life would improve to the extent that some residents' **behavioral patterns** began to transform in a positive, constructive direction, as promoted by changes on **living conditions** that, in turn, were to be induced by the submission of **project deliverables** specified in the intervention original design. Also according to our evaluation approach, verification of this hypothesis would depend on: (i) the degree to which the submission of *project deliverables* achieved or not its operational targets; (ii) the degree to which changes in resident's *living conditions* and *behavioral patterns* could be observed to be concomitant or subsequent to the project execution and the submission of those deliverables; and (iii) the degree to which observed changes on *living conditions* and *behavioral patterns* could, arguably, be attributed to the execution of the Project.
- 4.3 In general, concerning the all important question #1, an first general conclusion is that the Project attained or surpassed most of its targets for deliverables (*outputs*) and that they were submitted in

quantity, quality and opportunity adequate enough as to expect the induction of the desired transformations in the batey resident's living conditions and behavioral patterns; but that, unfortunately, the expected degree of such transformations (intermediate and final *outcomes*) was not achieved -as measured by the attainment of the Project's own outcome targets. A second general conclusion is that -despite the numerous gaps in the base-line values measured at the start of the operation- the analysis of those outcome metrics that provide an estimate of the ex ante situation do confirm changes in living conditions and behavioral patterns that can be time-associated with the execution of the Project. The great majority of those changes reveal improvements on the base line values. Finally, a third general conclusion is that, at least for those few available outcome indicators on which the quasi-experimental research design was fully applicable -based on comparison between treatment and control groups- the observed transformations in *living conditions* and *behavioral patterns* were indeed likely attributable to the BCDP's execution. With few exceptions, that will be discussed in particular, changes in the control group actually showed flat or backward trends in various degrees, implying an actual deterioration in living conditions and behavioral patterns, while those in the treatment group pointed at clear improvements.

1. What was expected from the Project

a. Objectives attained

- 4.4 Although the concept is not made explicit in any part of the documentation available, the researchers understand that the preponderant criterion used in setting out each annual target was a "cumulative" one: i.e. targets were in most cases established as increasing benchmarks, in which advances reached in periods past were added to the annual target each time around. So, to the extent that this was indeed the preponderant procedure, the one target set for the last year in each metric would represent most of the time an overall criterion for success of the development effort "as a whole", in each particular area of measurement. With this perspective in mind, the analysis provides mixed results. Concerning *total annual targets* -i. e. targets for both project *outputs* and *outcomes*- the data reveals that the Project reached or exceeded roughly two thirds (66%) of all its goals in phase I. Specifically 36% of total annual targets were reached on the mark and the other 28% was clearly surpassed in that phase. Given the normal working conditions in the Dominicans bateys, this assessment speaks of reasonable results for the Project; yet they get somewhat qualified when performance is judged against final year targets -a better criterion for overall operational success. The data reveals that 38% of final annual targets were not met in phase I of the Project.
- 4.5 Another facet of the Project's performance vis-a-vis its own objectives is manifest when target attainment is judge separately by type of development result. Indeed, a wide difference is revealed between the performance judged by the submission of *development outputs* -i.e. operational deliverables- and the performance judged by the achievement of *development outcomes* -i.e. changes in living conditions and behavioral habits in the target population. Project performance in terms of output delivery is strong. Based on the classification of result metrics into indicators of *outputs*, *intermediate outcomes* and *terminal outcomes*, 77% of annual *output targets* were met on the mark or exceeded by the Project in phase I. And when the final year output targets -again, a stricter criterion of success- is taken into consideration, the Project registers an even stronger performance with fully 83% of these

targets met or surpassed. Also, as can be verified in Table II.1 of Chapter II, the less representative Phase II of the Project registers somewhat lower but similar delivery record; with 70% of annual targets and 80% of project-end targets attained or surpassed. All this constitutes a *prima facie* appropriate basis to expect higher level development results in the batey's reality, as consequences of the Project's actions.

- 4.6 Now, it could be argued that those expected higher level development results would generally require time to emerge and take hold; especially those already bordering the frontier with preliminary impacts, such as initial changes in behavioral patterns of people. Yet, the BCDP executors did establish relatively early targets for such transformations, and when the fulfillment of those outcome targets is taken into consideration the Project's performance is not so impressive. Less than half of the annual *outcome* targets set out (48%) was actually met or surpassed in the bateys of phase I. Moreover, outcome inducement appears to have been even weaker when final year targets are taken into consideration. In fact, more than 70% of project-end outcome targets set by the Project were not attained. Otherwise, as per Table II.1 of Chapter II, the executors reported a much better record for phase II on this respect; with only 30% of annual outcome targets, and 21% of project-end outcome targets not being attained.

b. Changes confirmed

- 4.7 Only a fraction of the development result indicators defined included baseline measurement at the start of the BCDP; however the field data examined provides clear evidence of such changes that can be time-correlated with the execution of the project, in at least in the 10 indicators that were so-measured: all positive changes with the exception of only one. The documented changes spanned the gamut from health and wellness practices in adults and youngsters, to elementary school enrollment and graduation rates. The exception to the rule of advancement is the indicator of children enrolled in pre-school, which actually declined significantly in the SPM – Hato Mayor bateys.
- 4.8 There is no lack of examples of dramatic positive changes as well, in the SPM – Hato Mayor area: such as the one observed on the incidence of diarrhea in children under 5, which dropped precipitously at the same time of the project execution. Also large increases in the value of indicators, occurred concurrently with the intervention, have been document in relation to children health & nutrition practices; women reproductive health practices; and STI, HIV/AIDS prevention practices.

c. Results attributed

- 4.9 The detailed analysis presented in the present report allows some conclusions about the third, and more stringent success criterion applied in the present evaluation; namely: whether or not the development *results chain* (outputs, intermediate outcomes, terminal outcomes) observed in the wake of the BCDP can be statistically attributable to the workings of the Project itself, or may have resulted randomly from social evolution or happenstance. Even though no appropriate measures of development *impacts* -in the precise definition internationally accepted by the trade- could be worked out in the case of the present project, the field studies have provided enough data to examine the Project's *outputs*, *early outcomes* and *terminal outcomes*.

(i) Verifiable deliveries

4.10 Since the delivery of project outputs falls within the full control of executors, there can be no uncertainty on the attribution of such deliverables to the Project itself; and so, there is no need to test for statistical significance the differences encountered between treatment and control groups in the *output* tranche of the results chain. Yet verifying the output delivery to the treatment group is indispensable to provide grounds for attributing latter tranche results to the intervention of the Project. In the case of the BCDP, the evidence of delivery of expected outputs is clear and robust. The mentioned 83% rate of delivery on end-project output targets and the fact that in more than a fifth of the cases (22%) those operational targets were actually exceeded attest to the operational effectiveness of the Project. Finally the generalized positive perceptions of community leaders and field participants in the Project, as gathered in the study's qualitative interviews, confirm the view of the Project executors interviewed, and attest to the quality of the operational products submitted by the Project to benefit the intervened Batey communities.

4.11 The twelve bateys on San Pedro de Macorís – Hato Mayor intervened during phase I and II was by far the geographical area on which the full set of Project outputs was delivered in the most consistent manner in the areas of (i) health attention services; (ii) wellness promotion; (iii) refurbishing of school Infrastructure; (iv) enhancement of teacher competencies & tools; (v) revamping of sanitation infrastructure; and (vi) upgrading of water systems -in the main results chain- as well as additional benefits in the areas of housing repair and construction; emergency preparedness and family income generation. As discussed in detail on the report, the four bateys in the areas of Veron-Bávaro and Batoruco, included in phase II, also received a sub-set of those products. In the opinion of the present researchers, the submission of the mentioned operational products by the Project executors provided enough grounds to expect the eventual materialization of early and terminal outcomes as identified in the Project development hypothesis.

(ii) Verifiable early effects

4.12 Even though only a fraction of the identified early outcome indicators provided baseline data on both treatment and control groups, the ones that did generally yielded high values in the statistical significant test applied on the differences observed. And for those development outcomes where such statistical significance test was warranted, the obtained values constitute fairly conclusive evidence that the delivery of the Project outputs did induce changes in the living conditions of the intervened batey communities. This evidence includes indicators that, even if strictly speaking measure the delivery of outputs, could not reach noticeable values unless a degree of cooperation or change in attitudes is present on the part of the beneficiaries themselves. This is the case, for instance, of the observed outputs about: *women receiving pap smear tests or enrolled primary school students attending classes*. A variety of cultural or economic constraints, such as taboos about women's body or the frequent presence of child labor, traditionally inhibit or hamper these actions, in poor, uneducated communities. Therefore, the corresponding outputs could not really be delivered adequately, if those constraints did not somehow begin to crumble, and unless a modicum of disposition on the part of the population is present for the task at hand. This type of budding attitudinal change may signal a trend in the direction of more permanent changes in the behavioral patterns in the future. For this reason, we deem these results "intermediate outcomes" (initial or early development effects of the Project).

- 4.13 The documented intermediate outcomes observed -mainly in the intervened bateys of San Pedro de Macorís – Hato Mayor- point toward changes in the living conditions of residents such as: improved maternal, child, and reproductive health; STI and HIV/AIDS prevention; TB prevention and treatment; improved hygiene among batey residents; increased access to, and improved quality of, primary school education and extracurricular educational services; and improved water consumption and asepsis. Improvements in school and housing infrastructure, as well as other services, complemented the benefits.

(iii) Verifiable advanced effects

- 4.14 Even if the BCDP evaluation framework did not permit proper measurements of project impacts (i.e. contributions of the Project to long term, structural changes in the bateys' reality) the study has documented some advanced development outcomes already visible by the end of operations. These had the form of incipient changes in behavioral patterns in the target population which, if sustained, can be seen as precursor marks for more permanent transformations maybe coming in the future. These observed trends are time-associated and aligned with both the *development outputs* delivered and *proximal development outcomes* that appear induced by the BCDP, and do not seem to be explainable by pure chance.
- 4.15 The nascent changes in attitudes, behaviors and associated quality of living results that have been detected and can be attributable to the Project are substantiated and documented in the areas of: (i) health and wellness practices; (ii) education payoffs and (iii) sanitation habits of the targeted batey population.

2. What was not expected from the Project & the role of external factors

- 4.16 The differences found between base-line and end-line values, as discussed in Chapter II, documents improvements across the spectrum in the great majority the indicators so evaluable. The exception is the indicator on *children enrolled in pre-school*; whose numbers actually decreased notably in the targeted bateys during the years of the Project. In discussing with stakeholders this obviously negative unintended result, officials at *Save the Children* mentioned to have noticed a reduction over the years in the population of children of pre-school age in the bateys intervened: a possible cause for this indicator to have run contrary to the expected evolution.
- 4.17 Now, one possible consequence of the enhanced *reproductive health practices* clearly induced by the Project in the population might be a trend reduction of the birth rates in the targeted bateys. This might, in turn, explain a reduction in the age bracket associated with the indicator of pre-school enrollment; which, unfortunately, was defined in terms of the absolute number, and not the percentage, of pre-school age kids enrolled. A future ex-post evaluation may establish hard evidence of this trend and whether it had time enough to affect the behavior of the referred indicator. If this correlation were to be confirmed, we would be in the presence of a cross-effect factor, of the kind identified in the BCDP reconstructed *LogFrame* (see the *assumptions* column); factor that actually very well worked against intended results of the Project.¹²

¹² See the factor: "Unexpected cross-effects" in the *Assumptions* column (From components to purpose) of the retro-fitted BCDP LogFrame, in the *Note on the Evaluability of the Project*, page xii.

- 4.18 Another way in which the negative results on the *children enrolled in pre-school* can be viewed as self-inflicted by the Project's own operation is the fact, more generally associated with the BCDP, that the targets were often established separately for each year. Actually, the Project did not attain the stipulated annual target for this indicator on any year of execution. Now, if the above mentioned shrinking trend in the pre-school age population was visible to the executors, that fact should have prompted a revision of the corresponding annual targets, as it was admittedly done for other indicators; yet, executors maintained or raised such annual targets until the end of the Project.
- 4.19 As shown in Chapter II, the underachievement of Project targets was widespread -and also probably unnecessary because of the ambitious targets- concerning most project development outcomes. Now, that outcome targets were often not attained does not imply that the Project did not induce significant differences in the living conditions and behaviors of the batey population. Numerous examples of such positive changes are discussed in the text, such as those in the *reporting by care givers of key children health and nutrition practices*; or in the *identification by adults of STI and HIV/AIDS prevention practices*; or in the *hand-washing habits of kids before eating at school*. In all of these areas visible strides were made, but the Project did not quite reach its own stipulated final year targets for those indicators. So it can be concluded that BCDP executors and promoters have been indeed self-defeating in, sometimes, setting unrealistically high expectations, especially about attitudinal changes, only after a few years of induction and education; somewhat downplaying the fact that those transformations take some time to appear and with, perhaps, not enough regard for external factors possibly retarding results, such as the resistance that changes induced "from outside" always tend to generate in the targeted communities, given the prevailing culture and socio-economic reality.
- 4.20 The majority of findings discussed in Chapter III confirm expected differences in living conditions and nascent behavioral changes between the treatment and control groups that can be attributed to the Project. Frequent are the examples of flat or even negative trends in such conditions and behaviors observed in the Control Group, while the trend has proven positive in the Treatment Group. For example, the percentages of *babies exclusively breastfed* and of *caregivers reporting child health and nutrition practices*, remained largely constant in the Control Group, while increasing in the Treatment Group. Also there was a visible deterioration in the Control Group's percentages of *women applying reproductive health practices*; *adults identifying STD and HIV/AIDS prevention practices*; and various other measures of sanitation conditions and habits; the last ones likely associated with the lack of maintenance and deterioration of latrines, all in clear contrast with positive trends observed for the same living conditions and behaviors in the Treatment Group. However a clear case of an unintended outcome was the steep fall in the percentage of children receiving the DPT3 vaccine in the Control Group.
- 4.21 Indeed, the Project clear success in achieving the DPT3 vaccination of the infant population in the intervened bateys appears sadly correlated with a sharp decline of such service to infants in bateys not intervened by BCDP. Before the Project those kids used to receive DPT3 from other sources in about the same original proportion as that of the treatment Group, but somehow stopped receiving it in quite the same manner during the project execution. Although no relevant hard evidence is available, hypothetical explanations for this negative unintended result include the unfortunate possibility that government vaccination programs in had actually diminished their incidence in bateys nearby the Project targets; or that, given a fixed, and maybe insufficient official supply of vaccines in the area, the

quicker and more efficient use of them by the Project in the Treatment Group implied and actual reduction of vaccines available for those children in the Control Group. Whatever the explanation, that a specialized ex-post evaluation may eventually confirm, we are in the presence of a negative unintended result likely caused by an unexpected factor originating outside the control domain of the Project.

- 4.22 Other unintended outcomes have been observed but, this time, of a positive nature. These are the cases of results on *women receiving PAP smear tests* and the more advanced outcome of incidence of *diarrhea in children under-5 years old*. There may be a reasonable doubt that the positive results obtained in both indicators are really attributable to the Project, because the trend observable in both Treatment and Control groups shows an almost identical improvement in either indicator. In the case of the “hybrid” (output/outcome) indicator of *PAP smear test*, there can be no doubt that results in the Treatment Group are due the Project; because -as an *output*- the PAP smear test was actually given to women by the Project. In the case of the “pure” outcome indicator of *diarrhea in children under-5 years old*, a doubt is more justifiable on whether this change in health was really due to the workings of the Project.
- 4.23 Also, in discussing with stakeholders the results on the *PAP smear test* indicator, the possibility has been raised that an external factor might have at work to cause and explain the unexpected control group outcome; one that has also been identified in the Project’s reconstructed *LogFrame*.¹³ Though there is no hard evidence confirming it, this hypothetical cause of the identical trends in PAP smear test received in both control and treatment groups is that the two groups chosen at the start of the Project were in fact porous and not perfectly isolated from one another. If this has been the case, just a *word of mouth* could have spread the message across groups about the benefits of receiving the test for women who, although living in separate, distant communities, may visit and talk to each other regularly. Since the Project did not have the monopoly on PAP smear tests, such word of mouth may just have generated a positive “contagion effect” of sort on women’s behavior, regardless of the individuals’ locations or the source of the PAP smear test.
- 4.24 Treatment and Control groups selected at the beginning of the operation were supposed to be assigned randomly, and be impermeable enough from one another as to preclude influencing each others’ behavior directly or indirectly; lest extraneous, uncontrolled-for explanatory variables be introduced to the experiment. So the mutual permeability of the selected groups would be, arguably, a non-random, systematic measurement error in the present quasi-experimental study design. However, to be fair, the penetrability of social groups is a condition often nearly impossible to fully avoid in the case of normal, open human communities; and it is not one that you may necessarily want to avoid. The present researchers randomly chose the households to be interviewed in the final control survey, but they had to work within the bateys already chosen and fixed from the start of the BCDP as the Control Group, which may perfectly have been at least partially permeable with the Treatment Group.
- 4.25 Finally, social permeability could also explain similar unexpected “contagion effects” between communities, beyond that presumed in the *PAP smear test* case. Another possible case would be, for instance: new healthy habits, spread through example and word of mouth from the Treatment Group, ending up reducing the incidence of *diarrhea in small children* in the Control Group; although no firm

¹³ See the factor: “Social permeability of groups” in the *Assumptions* column (from components to purpose) of the BCDP retro-fitted LogFrame, in *Note on the Evaluability of the Project*, pag xii.

evidence supporting this circumstance is available either, in the present study. At any rate, and regardless of the real explanation for both results in the Control Group, they were not strictly intended by the Project, but clearly beneficial to the population involved.¹⁴

B. Effect on the participation, organization and mobilization of batey communities

- 4.26 No quantitative metrics, premises and approach to measure was defined at the inception of the Project for researching its effect on *community participation, organization* and *mobilization*; but the nature and content of views and opinions from project stakeholders provide useful qualitative evidence to judge Project effectiveness on these respects. From those views gathered and cross-checked by the present researchers in surveys and in-depth interviews with institutional and grassroots project stakeholders, this study concludes that the Project did systematically elicit collective self-management and mobilization; promote consensus building at the grassroots level about needs and required response; and followed a practical approach to discover ways to induce positive behavioral changes.
- 4.27 The above field approach is clearly time-associated with the current strength of grassroots organizations, the social cohesion, the awareness of development issues, and the active community initiative in solving identified needs, that has been verified by the surveys and interviews of randomly chosen community leaders and residents of intervened bateys. Especially effective to induce this organizational advances appears to have been the credibility-enhancing strategy of combining the training on subjects critical to people in the bateys -such as disease prevention, initial & vocational education, etc.- with the actual delivery of goods, services and tools also clearly beneficial to residents - in areas such as infrastructure for health, education and sanitation. The resulting organizational boost has been systematically corroborated by the interviewees, who speak fluently and knowingly not only of new vibrant grassroots organizations induced by Project's promotion activities, but also of the re-launching of old organizations that were previously dormant or dysfunctional. Generally people tend to perceive as high or very high the level of grassroots participation in the affairs of the community.
- 4.28 The fact that months after the closing of the Project the present study was able to confirm that grassroots organizations remained fairly involved in developing viable solutions to their self-identified problems speaks of an effect on social mobilization habits that went further than the spur of the moment when the Project activities were taking place. One message spread out by the Project that appears to have been particularly convincing for the inducement of enduring mobilization is the explicit goal made clear from the beginning that all goods and services provided by the intervention were to become of self-managed by the community for future sustainability. This appears to have resonated with communities whose past experience in mobilization seem to have been short-lived and only driven by political parties in times of elections. The visible community vigilance to keep in good working conditions the project revamped water and sanitation infrastructure, clearly contrasts in people's minds with their past experiences, when repairs and re-constructions made without grassroots participation, quickly deteriorated for lack of community attention and maintenance. Generally people tend to

¹⁴ Though not relevant to *development outcomes*, but to *deliverable outputs*, another external factor working against expected Project results, worth mentioning, was the labor conflicts between teachers and the national government which delayed the literacy training in the BCDP and led to project changes that ultimately affected the Math training.

perceive as high the level of success attained by the social mobilization in their communities, regardless of the relevant issue.

- 4.29 Another highly consensual effect of the Project, as confirmed by interviewed community residents and leaders is the positive influence of the BCDP in lessening internal community violence and conflict; especially in what concerns the resolution of the water conflict, all pervasive in the past. The resulting empowerment of groups and individuals at the grassroots level has been verified in the workings of disaster mitigation and response committees, local water boards and women's groups currently operating in the bateys under study. Residents spoke knowingly during interviews about the nature of such emergency plans and of specific instances in which they have been launched in practice: a hint of the real level of involvement of people at the grassroots level in emergency preparedness and response.

C. Effects in creating partnerships with key external actors

- 4.30 The study confirms that coordination between the Project and other institutional actors involved in the bateys did occur, especially when it was needed for the submission of Project deliverable in the areas of Health, Education and Sanitation. Cases in point were the launching of MISPAS Health reference systems; the provision of official contents for the training of batey school teachers in Literacy; and the controlled supply of vaccines for the Project's DPT3 immunization activities. A particularly close cooperation with government institutions was reported to take place concerning health services in the bateys of phase II in the micro-region of Verón-Bávaro. Also a case of successful partnering of the Project with the private sector was the coalition: USAID- Rotary International H2O Alianza to bring important improvements to the bateys in phase II. No concrete evidence, however, could be factually gathered, or confirmed by opinions from interviewees, that the Project was particularly successful in creating partnerships with key external institutional actors in the public sector.
- 4.31 Indeed, the residents did perceive and appreciate the useful coordination between the Project and public sector institutions in the particular instances in which it was necessary. However, an active partnering between the two as part of the enabling strategy for sustainability of efforts did not rank very high in people's perception. On the contrary, in giving generally low grading to some official involvement in the Project -as, for instance, to the dedication and quality of work of health promoters recognized or appointed officially by the Ministry of Health- interviewees hinted at a limited synergy achieved in the collaboration with the public sector.

V. Recommendations (for possible future efforts)

- 5.1 The all important possibility of lessons learning, and improving successive stages in a development process such as the one pursued by the BCDP operation, resides critically on defining clearly -and hopefully not over-ambitiously- the chain of results expected from the beginning; with the inclusion of clear markers and metrics for such a chain of results to be measured, monitored and evaluated. Despite the admitted effort made by designers to provide metrics and a monitoring system to the Program, researchers find that important improvements can be made in this area for future efforts.
- 5.2 As it is expounded in some detail in the *Note on Evaluability of the Project*, several methodological lapses were found in the Project's "Results Framework" and monitoring system. The original narrative in the elements of such framework did not conform to internationally normalized definitions, and the indicators proposed in original project documentation provided to researchers lacked in precision and

measurability. On the other hand, the effort that finally was made before the start of the Project to define measurable indicators was excessively prolific and imbalanced in favor of lower tranche results (*outputs*) and with no impact indicators properly defined. Finally there was a frequent lack of baseline measurements for comparison of progress, which made difficult to judge the general development effectiveness and efficiency of the Project;

- 5.3 Even if the metrics identified for monitoring and evaluation of the Project were overabundant, they fell short of being adequate to gauge the Project's quantitative performance as a development operation, interested in results beyond the submission of operational deliverables. Also targets appear not to have been set explicitly ex ante, at the moment of project inception, and were separately defined by year of execution. Now the targets established often proved to be too high for actual achievement. In developing the indicators system, the lesson already known from the best international practices appears to have been ignored that the adequacy of a project's metrics and follow up system is not measured by the abundance of indicators, or the ambition of targets, but by their quality, the feasibility of expectations and the value added for management feedback and development effectiveness evaluation. So, an argument could also be made that perhaps excessive energy went in this Project to measure detail operational nuances, and perhaps not enough attention to measure other higher order development results.

Recommendation # 1: Adopt a more streamlined, prudent, less ambitious and more measurable expected chain of results, for further stages of the effort.

Although extensions of the Program might allow the highest end results stipulated in the original *Log Frame* to fall into the realm of the "doable", it might be worthwhile for evaluation purposes to revisit the program expected results chain with a view to define not only more achievable development objectives given the time span, but also more measurable indicators of performance at these higher order objectives.

Recommendation # 2: Establish a more balanced and concise set of benchmark indicators, along with clearly defined means of verification, to the extent possible independent of management control.

Unambiguous and measurable indicators should be included in the monitoring system, especially for the upper tranches of the results chain that gauge development outcomes and impacts. Beyond the realm of requisite management control (inputs and outputs) means of verification should establish clear responsibilities and cost of measurement, independent of Program interference. One or two high pertinence and quality indicators should be sufficient per each element of the results chain, including the proximal or intermediate results envisaged.

Recommendation # 3: Measure proper baseline values for each performance indicator in future stages of the effort.

Each performance indicator should have a target value stipulated and a baseline value measured, as well as proper specifications of quality, quantity and opportunity consistent for both the target value and the baseline value. Past cumulative achievements and experience should facilitate the definition of indicators and measurement of baseline values. Control groups might be considered for use in later evaluation and comparison, in cases of hard to measure baseline values. Benchmark (target) indicators, means of verifications and baseline at the level of terminal impacts should be available for purpose of ex-post evaluation of the Program

Recommendation # 4: Include in the design of further stages of the effort an analysis of effectiveness risk factors, to be used in management and evaluation.

Major threats possibly impinging on the Program's result chain should be ex-ante identified as thoroughly as possible. The idea is to provide grounds, as practical as possible, to inform a risk management strategy for each element of the results chain.

Technical references

- A. Weis, C.H. (1998) "Evaluation: Methods for studying programs and policies." Prentice Hall, NJ.
- B. Campbell, D.T, & Stanley, J.C. (1963). *Experimental and Quasi-experimental Design for Research*. Chicago, Rand McNally.
- C. Cohen, J. (1992). "A power primer". *Psychological Bulletin*. 112 (1): 155–159. doi:10.1037/0033-2909.112.1.155. PMID 19565683).
- D. Faul, F. (2012) GPower Version 3.1.4,
<http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register>
- E. Preacher, K. J. (2001, April). *Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence* [Computer software]. Available from [%http://quantpsy.org.%](http://quantpsy.org)

UNITED STATES AGENCY FOR INTERNATIONAL DEVELOPMENT

BATEY COMMUNITY DEVELOPMENT PROJECT

DEVELOPMENT EFFECTIVENESS EVALUATION STUDY
(ANNEXES)

SANTO DOMINGO, D.R.

SEPTEMBER 2013

ANNEX I: ORIGINAL PROJECT THEORY

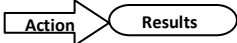
4. Results Framework

Goal: Selected pilot bateyes in San Pedro de Macoris and Hato Mayor have community based and led programs in place that markedly improve residents' quality of life		
Strategic Objective: Residents of the selected pilot bateyes improve their standard of living through community based programs in health and nutrition, education, infrastructure and income generation, and program learning is documented to enable replication of the model		
Result 1: Health & Nutrition	Result 2: Basic Education	Result 3: Infrastructure
Improved maternal child care, reproductive health, HIV/AIDS prevention, nutrition and hygiene among batey residents.	Improved access to and quality of primary school education and extracurricular educational services.	Improved water, sanitation, housing, schools, and health facilities including at least one emergency-safe communal structure per community.
Strategic Interventions		
1.1 Improved access to quality basic health services for maternal and child care through community providers	2.1 Increased access to basic education through school infrastructure improvements and equipment and/or materials	3.1 Improved community access to sustainable (fee for service) water and sanitation infrastructure
1.2 Improved knowledge, attitudes and skills regarding key maternal-child health and nutrition practices and services, including family planning	2.2 Improved quality of education through enhanced teaching quality	3.2 Access to at least one emergency-safe communal structure per community (school or other community building)
1.3 Improved understanding of safe sex practices including prevention of STI and HIV/AIDS	2.3 Improved enrolment and attendance of children in grades 1-4	3.3 Improved housing, school/classroom infrastructure, and community health facilities
1.4 Improved hygiene practices including hand-washing, use of latrines, safe food preparation, consumption of potable water, and community environmental health	2.4 Children and youth have access to and use extracurricular activities and groups which complement formal curriculum	3.4 Emergency preparedness plan for each batey developed and promoted with involvement of all community members, including children, youth, women and community leaders
Enabling Strategies		
Participatory Community Mobilization Develop or strengthen mutually reinforcing networks, including children's clubs, youth groups, women's groups, water management groups, <i>Asociación de Madres, Padres, y Amigos de las Escuelas</i> (AMPAES), and infrastructure and emergency preparedness committees at the community level in each batey to engage, motivate, and support community members to sustainably work together to shape positive physical and behavior changes in their community.		
Income Generation Improved access to means of income generation and improved household nutrition through micro-enterprise development, small scale agriculture, and small animal husbandry.		

Source: Annex I- BCDP Program Description

ANNEX II: NOTE ON THE EVALUABILITY OF THE PROJECT

A. Conceptual approach and constrains

1. As defined in the research proposal, researchers were to evaluate this project by first assessing whether it achieved or not its own stipulated objectives (*achievement criterion*). Other unintended or indirect project consequences, as well as the role of any external factor at play in shaping the actual results were also to be studied, but only second to judging effectiveness against the project's original intent. Also the study was to find out: (i) if observed changes in the targeted communities are concomitant with the Project's execution (*association criterion*); and (ii) if the observed changes were statistically attributable to the Project (*attribution criterion*). All this brings to the fore the issue of *evaluability at entry* of the project, i.e.: the extent to which the operation's performance can be judged against a pre-ordained set of measurable objectives and other metrics conditions established in the project's documentation.
2. The modern conceptual approach to judging development effectiveness requires that projects identify ex ante the *results chain* they pursue with their execution, as well as the set of associated performance metrics, requirements for attribution and envisioned ineffectiveness risk factors.¹⁵ Based on the definitions by the *Development Assistance Committee of the OECD*, our operational view of the result chain is depicted in figure A1¹⁶. If a development program can be succinctly understood as the binomial expression:  then the "results" member can be thought to include the sequential set of *outputs*, *outcomes* and *impacts* expected from the deployment of the program's inputs; set which normally is laid out in a program's so called Logical Framework or *log frame*. For purposes of the present research we define the elements of the relevant results chain as follows:

- (i) **Outputs:** First tangible results. Clear throughput coming directly off project activities. Outputs are usually equivalent to the so-called program "deliverables", upon which management has maximum direct control.

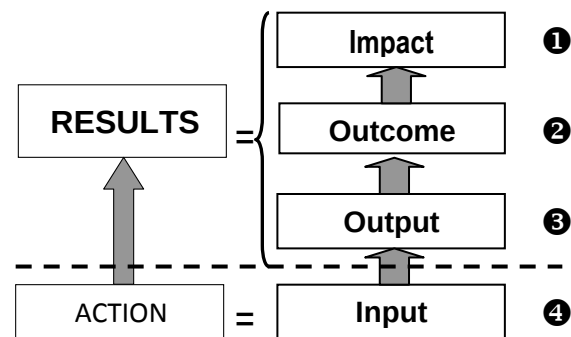


Figure A1: The expected Results Chain

- (ii) **Outcomes:** Second tier results. Immediate changes taking place in the program's target reality, imputable to a project's output delivery, or output's onset and workings, at the end of execution. Project managers normally have no direct control on this category of results, but outcomes form part of the causal hypothesis explicit or implicit in most project designs. From an ex post viewpoint, outcomes may be directly or indirectly attributable to outputs and deemed positive or negative, intended or not intended.

¹⁵ See proceedings of the *International Conference on Financing for Development*, Monterrey, México, 2002; and the *Marrakech International Roundtable on Results*, 2004

¹⁶ See OECD DAC Working Party's *Glossary of Key Terms in Evaluation and Results Based Management*

- (iii) Impacts: Third tier results. Changes attributable to the project, taking place over a longer time span after execution has finished, and/or in a wider, more complex and farther reaching context surrounding the project's target reality. Control over this category of results is even more problematic and attribution more difficult to establish or argue; usually requiring from the outset the establishments of control groups and experimental designs and settings. Yet, impacts also form part of the causal hypothesis explicit or implicit in most project designs. From an ex post viewpoint, impacts may be directly or indirectly attributable to outcomes and deemed positive or negative, intended or not intended.
 - (iv) Intermediate results: Earlier outputs, outcomes or impacts, attributable to the project. Intermediate outcomes can possibly occur while the project is still under execution. From an ex post viewpoint, intermediate results may be directly or indirectly attributable to the related elements in the results chain and dimmed positive or negative, intended or not intended. In "ex ante" *log frame* parlance, impacts correspond to the professed *goal* of a project; outcomes relate to its intended *purpose*; and outputs are directly associated with the program's *components*.
3. Although by its scope of work the present evaluation's purported focus is the high-end portion of the results chain (possible *impacts*) the actual research does not ignore precursor elements in the results chain, for two main reasons. First, the exclusive emphasis on establishing impacts without verifying in the same breath the concomitant occurrence of precursor outputs and outcomes would leave the final argument without proper grounds for attribution (i.e. to be able to impute the former to the latter, by at least arguing time correlation thereupon). Secondly, given the relatively short time elapsed since this project's start, impacts -as strictly defined- may very well have not had time to fully mature yet, while other important elements of the results chain -especially outcomes- might arguably be more feasible and currently visible. The present evaluation, therefore, studies the project's whole result chain to the extent that is evaluable at this point in time. Within the approach outlaid above, the researchers' first order of business has been to identify clearly the project's results chain as it was envisaged in the original documentation. We also tried to pinpoint established benchmark indicators, verifiers and base-line indicators needed to gage project performance vis-a-vis its intended objectives; as well as other metrics requirements for attribution of results.

B. Is this Project's expected results chain clearly identifiable ex-ante and evaluable?

4. In order to examine the Project evaluability, following the methodology discussed in the preceding section, the researchers investigated if the project's expected development *results chain* was clear and explicit at the inception, and whether the set of indicators, verifiers and base-line information provided at the start of the study was complete and consistent with technically acceptable practices to gage a project effectiveness vis-a-vis its intended objectives. On this score the review renders mixed results, as the requirements for evaluability at entry defined above appear only partially fulfilled in the Project's documentation. Designers clearly made an effort at the inception to define expected results. They also identified metrics to gage such results. Yet, relevant specifications were somewhat lacking in methodological accuracy, and other critical deviations occurred as well with respect to standard practices.

5. Three areas of concerns were identified. First, there is no clear conceptual correspondence between what is stipulated in the project's *Results Framework* in page 15 of the "Program description" and the standard notion of *Results Chain* as internationally accepted today. Second, a lack of clear correspondence was also found between: (i) the indicators stipulated in pages 5 through 7 of the S.O.W., (ii) the ones stipulated in the Program Description, and (iii) and the "final" list of indicators provided by Executor, for use in the base-line studies. Third, the majority of indicators in lists (i) and (ii) were stipulated as "pure" -non directional- rather than *bench-mark indicators*; precluding a standard "normative" evaluation. Therefore, the Project's evaluation premises needed to be reviewed to ensure they were aligned with the researchers' technical proposal, approved for the present study.
6. Conceptual limitations of the Project's original "Results Framework". For the case of the BCDP, the *means-to-end* logic based on accepted definitions of *results* and *results chain* (INPUTS-OUTPUTS-OUTCOMES-IMPACTS) that should be depicted in *LogFrame* matrices, can be understood as illustrated in Table B1. This table presents micro-level examples, in the areas of education and health, of results/indicators actually identified in the Project documentation.¹⁷

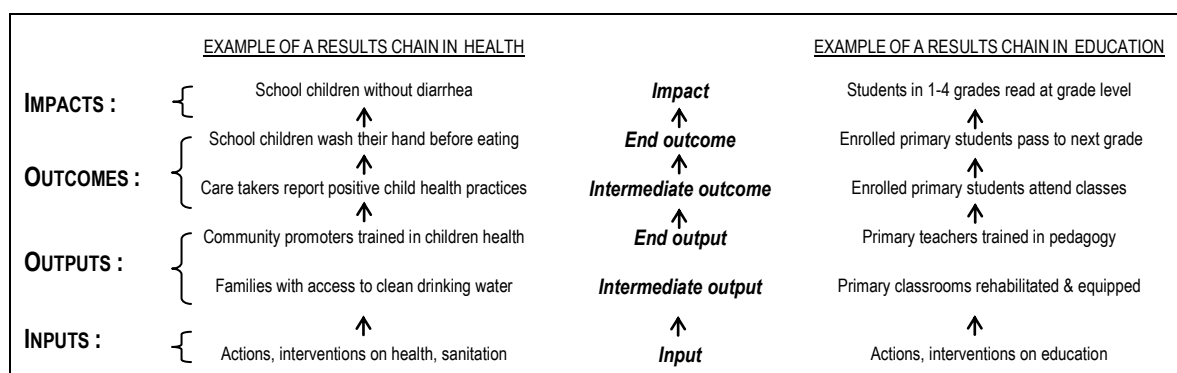


Table B1: Micro-level examples of Results Chains in the BCDP

7. As it was discussed in the previous section, the standard concept requires that each and all of the elements of the chain be identified as different, separate facts; which could then be correlated to one another and a *cause-effect* relation (project theory) may be argued, or disproved, between them. Unfortunately, the BCDP's *Results Framework* as stated in the original documentation seldom presents results that can be clearly understood as different, or measured separately in each of its different rows. Often, the elements presented in the "Results" row -whether these correspond to expected *outcomes* or *outputs* is not clear- appears to be just re-phrasings or summaries of the same elements included below in the "Interventions" row -whether these correspond to expected *outputs* or *inputs* is also not clear either. The same occurs with the stated "Strategic Objective" (presumably an expected *outcome*) which appears to merely re-phrase and summarize the same elements presented in the "Results" row below. Adding to the lack of a clear-cut distinction between elements in the *Results Framework*, sometimes even the same language is used to describe allegedly different elements. This is the case of the phrase

¹⁷ In results chains of a more "macro" nature, outcomes and impacts may refer to more general, farther-reaching transformation results. The present examples are intended only to illustrate the logic behind the concepts.

“improved access and quality” used to describe both “results” and “interventions” in the area of Basic Education; or the phrases “standard of living” and “quality of life” -arguably synonyms- used both for the “Strategic Objective” and the “Goal” (presumably an expected *impact*) of the Project.

8. The “intervention” row alone provides a somewhat more concise and, therefore, measurable -albeit implicit- definition of the Project expected results, than any other component of the *Results Framework*. Yet, the intervention row arguably mixes up results of different levels in the same category. For instance: *improved schools infrastructure and equipment* (an output) with *improved quality of education* (an outcome); or *improved health knowledge and skills*, (an early, or intermediate, outcome) and *improved hygiene practices* (a later outcome). Due to this level of ambiguity in the language used, the present researchers did not use the original version of the *Results Framework* as basis for the present evaluation study. Instead, we searched for other elements in the documentation that would allow a better alignment of definitions with standard *LogFrame* concepts and practices. Fortunately, the situation just discussed improved in the section of the project documentation dealing directly with indicators, both in the *Program Description* and in the *Study S.O.W.* This afforded an indirect way of “discovering” expected development results, because the effort to specify results metrics provided additional clarity as to what precisely was what the Project wanted to achieve. Accordingly, this yielded a basis for a reconstructed *Results Framework* better focused on measurement. Yet, a number of limitations also affect the quality of metrics thereby identified.

C. Are ex-ante metrics requirements for evaluation present in this project?

9. Conceptual limitations of the indicators originally identified. The main difficulties found with the Project metrics are: (i) inconsistencies between the indicators illustrated in the Study S.O.W. and the ones stipulated in the Program Description; (ii) bias toward outputs and against outcomes and impacts in the metrics mentioned in the documentation; (iii) inconsistencies between the metrics identified in the project documentation and the ones identified by executors for the base-line studies. The indicators illustrated in the *Study S.O.W.* are 20 while the ones stipulated in the *Program Description* are 28. The two lists are also largely inconsistent, for they have in common only 9 indicators. The fact that the *Study S.O.W.* list is shorter is not of major concern, because its purpose was only to “illustrate” some of the indicators defined in the latter -and, maybe, the ones used later in the base-line studies as well. The inconsistency that does matter is that the “illustrative” list includes indicators that were not included in the *Program Description* or, for that matter, eventually used in the actual base-line studies.
10. Concerning the nature of results to be measured by the identified indicators, only two (2) of the list of twenty indicators illustrated in the *Study S.O.W.* can be unambiguously said to refer to an *outcome* or a higher level development result; while the majority (11) clearly measures only project *outputs*. The seven (7) remaining indicators are of a more ambiguous nature. The proportion of implicit higher level expected results increases in the list of the *Program Description*, with only 10 of the 28 indicators clearly referring to early *outputs*; therefore, that list provides more balanced mix of expected results upon which to base a development effectiveness evaluation. However, the rest are of ambiguous nature, between *late outputs* and *early outcomes* (designated “p/o” in our tables) Also, this latter list still does not fully conform to the Final List provided by Executors for the base-line studies.

11. The number of metrics in the *Final List* runs up to 41 indicators, and this list is not fully consistent either with that in the *Study S.O.W.* list, or in the *Program Description*. Of the 41 indicators, only 9 were also listed in the S.O.W. and 28, in the *Program Description*. This reveals that at the moment of compiling the *Final List* at least 13 new indicators were added beyond those stipulated in the other lists. Tables C1 and C2 demonstrate these two facts; namely that: (i) despite the profuse number of indicators mentioned at one time or another in the Project only a small fraction (18,5%) was present in all instances; and (ii) a significant proportion of indicators (24%) were actually coined up after the design of the project.

<u>INDICATORS ONLY IN THE STUDY S.O.W.</u> - # of health facilities refurbished/equipped/rehabilitated (p) - # of antenatal care visits by skilled providers (p) - # of children reached by USG-supported programs (p) - # of people benefitting from community-based public health and nutrition messages (p/o) - # of batey residents demonstrating improvement in health and nutrition behaviors (o) - # of people that have seen or heard a specific USG supported FP/RH message (p/o) - # of people trained in FP/RH with USG funds (p) - # of USG programs interventions giving counseling and/or community awareness activities to respond to/reduce rates of gender based violence (p) - Model batey committees are operating and maintaining water and sanitation systems with a "pay for service system" (p/o)
<u>INDICATORS ONLY IN THE PROJECT DESCRIPTION</u> - % CHP following appropriate community management protocols (p) - Increased % of children 12-23 months fully immunized (p) - Increased # of batey community with improved environmental health (o) - Increased # of teachers applying innovative teaching methods (o)
<u>INDICATORS ONLY IN THE "FINAL LIST"</u> - % of students in single grade primary schools in 1-4 grade who read at grade level (o) - # of community health promoters trained in TB management (p) - # of community health promoters trained in maternal-new born health (p) - # of community health promoters trained in child health and nutrition (p) - # of community health promoters trained in reproductive health/family planning (p) - # of children under 10 years of age who are de-wormed by USG-funded programs (p) - # of bateys with functional MISPAS reference systems (p) - # of adults tested for TB (p) - # of children enrolled in pre-school (p/o) - # of children and youth (6-14 years old) enrolled on USG supported extra-curricular programs (p/o) - # of youth and adults participating in USG supported vocational or informal education programs (p/o) - # of bateys with updated emergency plan (p) - # of residents receiving support in income generating activities (p)

Table C1: Indicators mentioned only in one list; with (o) standing for outcomes, and (p) standing for products or outputs

12. The above does not argue well for the existence of an unambiguous, clear-cut framework for evaluation at the inception of the Project. On the contrary, the definition of expected results and metrics appears to have been done in starts and fits, perhaps reflecting the "moving-target" pattern that appears to have been a general feature of the Project.¹⁸ In response, the researchers endeavored to remedy the situation by making the best use of the practical possibilities at hand. Specifically, we looked for a maximum possible alignment of the data provided by previous studies with the metrics originally envisaged. So, with this "retrofit" alignment in mind, the evaluation used metrics that were common both to the *Study S.O.W.*, to the *Program Description* and to the *Final list* identified by Executors, as proxy descriptors of the results implicitly expected by the Project. This list of indicators is reflected in the portion colored in shades of green in Tables B2 and B3; which covers the 41 indicators indicated in the

¹⁸ Since all metrics were not defined "ex ante" -some even as late as "phase II"- often no measure of the situation before the Project was include, which contributed to the scarcity of baselines against which to judge performance.

Final List. To the extent that the base-line studies provided field results for these indicators, the evaluation could be based on a Results Framework that allow a clearer measurable comparison between the situation of the bateys before and after the Project. Also notice that only a third (11) of the selected indicators corresponds clearly to expected *outcomes* -“o” in parenthesis in the table. The rest (30) corresponds to expected *outputs* -“p”, for products and, in parenthesis in the table- or, ambiguously, to *late outputs* and/or *early outcomes* -“p/o” in parenthesis. None of the indicators does actually qualify as a measurement of *impacts*, according to our working definition.

INDICATORS IN THE STUDY S.O.W. & THE FINAL LIST - % of children under 5 with diarrhea in the last two weeks (o) - % of children under 12 months of age who have received DPT3 (p) - # of administration and education officials trained (p)
INDICATORS IN THE PROGRAM DESCRIPTION & THE FINAL LIST - # of classrooms with improved didactic materials (p) - # number of community health promoters equipped (p) - % of care givers who reported at least three key positive child health and nutrition practices (o) - % of infants under 6 months of age exclusively breastfed within the past 24 hours (o) - % of women between the ages of 15 and 49 receiving at least one PAP smear test in the past 12 months (p/o) - % of women between the ages of 15 and 49 who can report at least two key reproductive health practices (o) - % of heads of household who identify at least two key STI and HIV/AIDS prevention practices (o) - % of primary school children with improved hygiene practices (who wash their hands after using the school latrine/bathroom) (o) - % of primary school children with improved hygiene practices (who wash their hands before eating the school breakfast/snack/lunch) (o) - % of enrolled primary schools students attending (o) - % of students enrolled in 4th grade that pass 4th grade (o) - # of families benefitting from improved community water systems (p) - # of families with improved housing (p) - # of bateys with access to an emergency-safe structure (p) - # of batey schools with updated emergency plan (p)
INDICATORS IN ALL THREE LISTINGS - # of people benefitting from USG supported health services (p) - # of children under 5 years of age who received vitamin A from USG supported programs (p/o) - # of classrooms rehabilitated and equipped (p) - # of teachers from participating batey schools trained in literacy (p) - # of teachers from participating batey schools trained in math (p) - # of teachers from participating batey schools trained in computer use (p) - # of teachers from participating batey schools trained in school health and nutrition (p) - # of teachers from participating batey schools trained in other pedagogical improvements/practices (p) - # of families with access to improved sanitation services (p) - # of families with improved access to clean drinking water (p/o)

Table C2: Indicators mentioned in two or more lists; with (o) standing for outcomes, and (p) standing for products or outputs

D. Are ex-ante requirements for measuring and attributing changes met by this project?

13. Measuring limitations of the metrics (indicators) identified. As the researchers pointed out in their technical proposal, when compared with the use of so-called *benchmark indicators*, the use of *pure indicators* (also called non directional, i.e. without a specified target value) implies a completely different perspective in the way a Project's performance is to be judged.¹⁹ If at all possible, the

¹⁹ In the standard *log frame* model, metrics are not understood as just “indicators” in the common parlance sense of the word. The model requires “benchmark indicators”; i.e., metrics that, beside a precise formula for measuring changes in the relevant variables, also include a “standard” (target or reference value) against which to judge changes actually achieved and in comparison with a starting value (base-line). This approach directly associates metrics with Project objectives and provides a clearer-cut gage for the “desirability” of actual results. For instance, as opposed to the indicator: % of children under 5 years old receiving vitamin A, the benchmark indicator: % of children under 5 years old receiving vitamin A to increase from 15% to 80% in 5 years allows project/policy managers to judge more precisely if an increase actually achieved was what was really desired.

researchers intended to use benchmark indicators in their evaluation study, not only because they comply better with current best development evaluation practices but, because by providing clear directionality to interventions, the use of benchmark indicators aligns also better with the current USAID Agency Evaluation Policy that, for purposes of accountability ... *requires comparing performance to ex ante commitments and targets...*²⁰. Unfortunately, the BCDP documentation originally handed to the researchers provided very little in the way of indicator targets. This, plus important gaps also found later in terms of base-line values for indicators, left evaluators with little to work with in terms of measuring development changes (outcomes, impacts) that might be associated with the Project.

14. Now, in spite of the above, the researchers recognize that there can be valid reasons to choose non directional indicators; not the least of which is that specific *outcomes* and *impacts* targets may be hard to establish, especially in the absence of enough proven theory or experience in particular interventions. This appears to be the case of the BCDP original documentation submitted to the present researchers, where the metrics identified initially for the Project were almost always “pure”, non-directional indicators, with the exception of a handful of metrics enumerated on the *Study S.O.W.* as “Overall BCDP project results” (see p. 5 of said document, where five of the six indicators mentioned include target values). Also, the phrasing of the indicators in the list of the *Program Description* “evokes directionality” to the extent that it starts with expressions such as “increased %”, “increased #” etc.; yet the indicators do not quite include a *target value* as such upon which to base a truly normative judging of achievements -that is, beyond the generally unacceptable idea that any improvement whatever, and however minuscule, were to be considered enough to deem any effort successful. Therefore, while recognizing the very few and far apart exceptions mentioned, if it were for the prevailing nature of the indicators included in the *terms of reference* for the present study, we would have been restricted to produce an evaluation focused on: (i) providing a non-normative judgment on performance (i.e.: not based on comparison to ex-ante stipulated target values); and (ii) judging only the incidence and quantum of Project results, not necessarily their direction.
15. Very fortunately, this restriction was lessened by the fact that, eventually before the start of the project, the executors did establish targets for each indicator in the *Final List*, as reported in their *Annual Project Report* (APR) of which researchers learned later. These targets appeared to have been “moving” -i.e. often defined anew at the beginning of each execution year. So, strictly speaking, they would provide no ex ante criterion to judge performance. Yet, the annual targets were often defined as cumulative benchmarks -that is: in the value for each new period they included the level of results already achieved in the previous period; and so, such targets do afford some measure of comparability between expected and actual results for the project as a whole. Accordingly, the researchers dedicate a full section of the present study to report on this comparison between annual targets and results reported by the Project’s Executors in their Annual Performance Report for 2012 (APR2012) which, as the final operational report, submits the full set of expectations and achievements for the whole project since 2008 through 2012. This comparative analysis, including the comparison between base-line and end-line values when

²⁰ See *Purposes of Evaluation* USAID Evaluation Policy, www.usaid.gov/evaluation/USAIDEvaluationPolicy.pdf, p. 2

possible, is intended to register for the record the information submitted by the management in charge of project execution in their own reporting. A separate, more substantial section is dedicated to review and discuss the field data collected in empirical studies, both previous and our own field surveys, that validates, or not, the Executors claims and constitute the independent evidence on the Project development effectiveness.

16. Limited scope of attribution of results afforded by the metrics (indicators) identified. Finally, the fact is noteworthy that such copious effort made by designers and executors to identify indicators resulted in so very short a list of metrics focused beyond the mere outputs of the project. Verifying the *development outputs* (products of a development operation) is, of course, important not only for purposes of precisely recording the actual submission of expected deliverables by the executing organization, but also to provide an important caveat for the attribution argument, namely: the temporal sequencing between the delivery of products and the ensuing *development outcomes* and *impacts* that might be observable in the field. However, the gaps that are truly worth investigating through the type of quasi-experimental design established for the present study are those in later-tranche results -outcomes, impacts- that might be found between *treatment* and *control* groups; and there is little comparative value added in such design for the mere reporting of *outputs*.
17. A short reflection on the underlying issues should make this clear. Figure D1 illustrates the general hypothesis that -in terms of differences between treatment and control groups- the present study seeks to prove; i.e. that the intervention would induce visible and sustainable changes (development *outcomes* and *impacts*) in the population intervened (line between points To_1 and To_n in the figure) as compared with changes in the same indicator observed in population not intervened (line between points Co_1 and Co_n).²¹ Such comparison provides useful information not only in terms of the vertical differences or gap between any of the inferred points of the two lines (v.g. dashed vertical segment Co_i-To_i) but also in terms of the relative slope of each of the compared lines (inferred trend) for purposes of estimating or projecting the comparative merits/efficiency of intervening in a particular situation, vis-a-vis the no-intervention alternative in the future.
18. All elements of the two segments in figure D1 (To_1-To_n , Co_1-Co_n) appear as legitimate objects of investigation through quasi-experimental research designs, and could only be established through an empirical field research. A particularly important challenge of the present field research is to establish that any observed final gap between results of treatment and control groups could not be explained away by

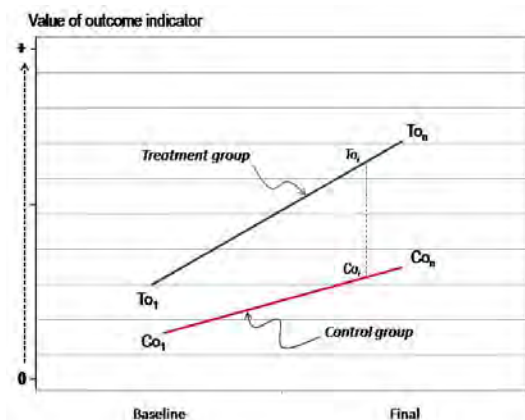


Figure D1: Example of change hypothesis, outcomes.

²¹ It is worth to note that, for this study, the field data corresponding to points To_1 , To_n and Co_1 (*base-line* and *project-end* values for Treatment Groups, and *base-line* values for Control Groups) represent data already researched on randomly assigned groups by the previous studies hired by the BCDP executors, and to be given as inputs to the present researchers, along with those randomly assigned groups. The data corresponding to point Co_n (*project-end* values for the Control Group) are the responsibility of the present study, as well as the overall comparison and analysis based on the randomization done at the start.

mere chance; reason why the study includes tests for *statistical significance* of the found differences (i.e. tests to establish confidence that such observed differences can actually be attributed to the Project). In the present study, to analyze the differences of the type depicted in figure D1, the researchers have applied to empirical data obtained the so-called χ^2 (*Chi-square*) test; which provides the adequate statistical parameters (**p**, **power** and **w**) to avoid confounding mere happenstance with actual correlation between results -or the opposite mistake of attributing correlation when there is in fact none- thus adding robustness and precision to the attribution argument.²²

19. A completely different situation is found in the case of indicators of pure delivery of project *outputs*, whose working hypothesis is illustrated in Figure D2. The only point to investigate, or empirically confirm, in this hypothesis is the one corresponding to Tp_n (*project-end* values for the Treatment Group). Since the indicators refer to products delivered only by the Project itself, initial values for both treatment and control groups are predictably known to be 0, as it is also predictably 0 the projected value and the slope of the line corresponding to the Control Group. The gap in this case may be fully attributable to the actual delivery (or not) of the expected products, with no possible confounding factors that may explain differences found. Since no attribution mistakes are possible, there is no need here for statistical significance tests or any other of the precisions afforded by the quasi-experimental study design. Unfortunately, the fact that only a few of indicators defined at the start corresponds to latter-tranche results, plus the fact that some of those few were not measured at the base-line for the control groups either, leaves the present evaluation with a very small number of indicators to make useful comparisons.

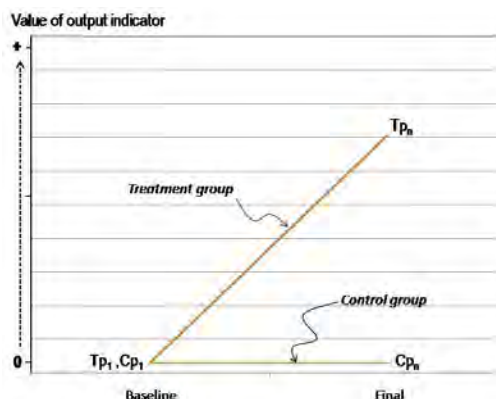


Figure D2: Example of change hypothesis, outputs.

20. Within the above discussed limitations, and looking to make the most of the practical possibilities at hand, the researchers have reconstructed the BCDP Results Framework in the way presented on the following section, in order to adequately gage the BCDP's performance with the available indicators.

E. A reconstructed Project results chain and its metrics

21. Although some indicators identified in the *Final List* for the base-line studies were thought of after the Project original design, they still were conceived at the start of execution. So they were to measure the "ex-ante" bateys' situation, allowing later comparisons with the situation after the Project intervention; which is one expressed goal of the present study. Therefore, the expected results chain implicit in such

²² Specifically, the **p** parameter measures the probability of attributing results to an intervention when, in fact, they are due to chance (type α error). So, the lower the value is of this parameter the higher the confidence is in attributing results to a project. In turn, the **Power** parameter measures probability of not making the opposite mistake: i.e. not recognizing attribution of results, when they can indeed be attributed to the project (type β error). Therefore confidence in attribution increases, the higher the value of this parameter is. The parameter **w** (Effect Size) estimates how strong the influence of the intervention is presumed on the results observed. Therefore confidence in attribution increases, the higher the value of this parameter is.

indicators can be used to rethink the “development logic” of the Project while allowing, at the same time, clear measurements. Based on this reasoning -and except for the tranche of *impacts* which, as explained above, could not be discovered in the metrics- figures E1, E2 and E3 below presents the results chains for each area of the BCDP intervention, reconstructed for the present evaluation.

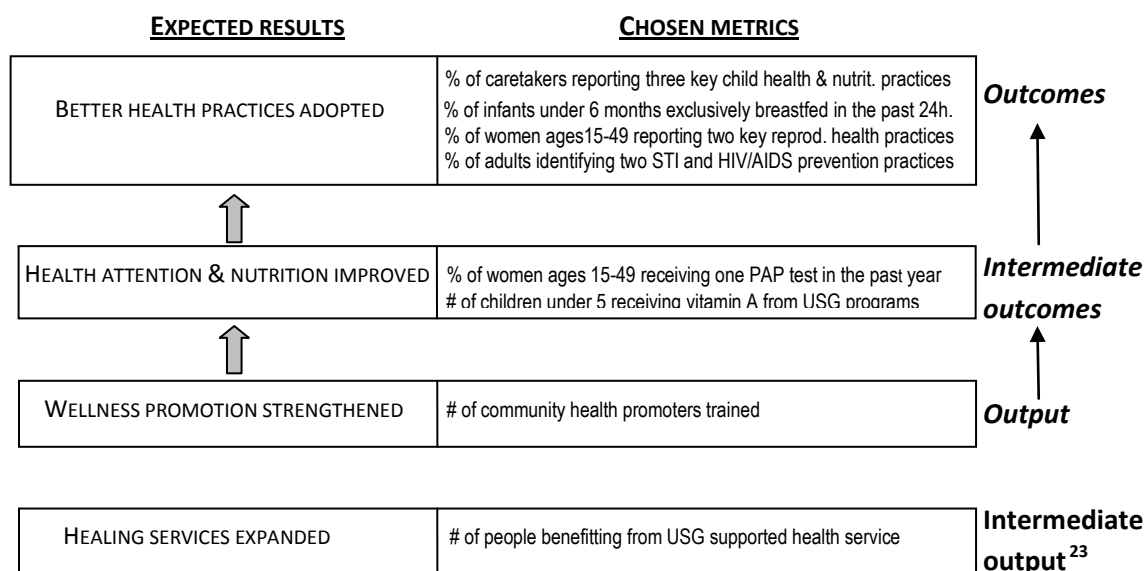


Figure E1: Example of BCDP reconstructed Results Chain in Health

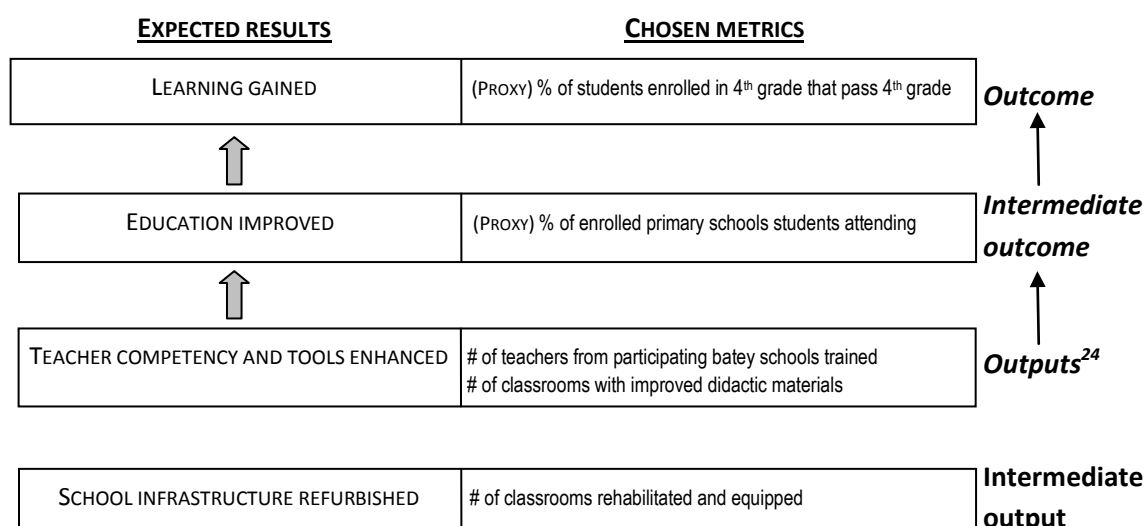


Figure E2: Example of BCDP reconstructed Results Chain in Education

²³ *Inputs*, as well as other elements, are also sometimes included in the definition of a generalized *Results Chain* (See, for instance, OECD, DAC Working Party's *Glossary of Key Terms in Evaluation and Results Based Management*, page 33). However, not being “results” of the intervention in any strict sense of the word, *inputs* will not be included among the subjects of the present research as a relevant part of the Results Chain.

²⁴ The indicator of teacher training for this output would include measuring the preparation of teachers in the areas of (i) literacy; (ii) mathematics; (iii) computer use; (iv) school health and nutrition; and (v) other pedagogical improvements/practices; all of these measurements corresponding to distinct and separate metrics.

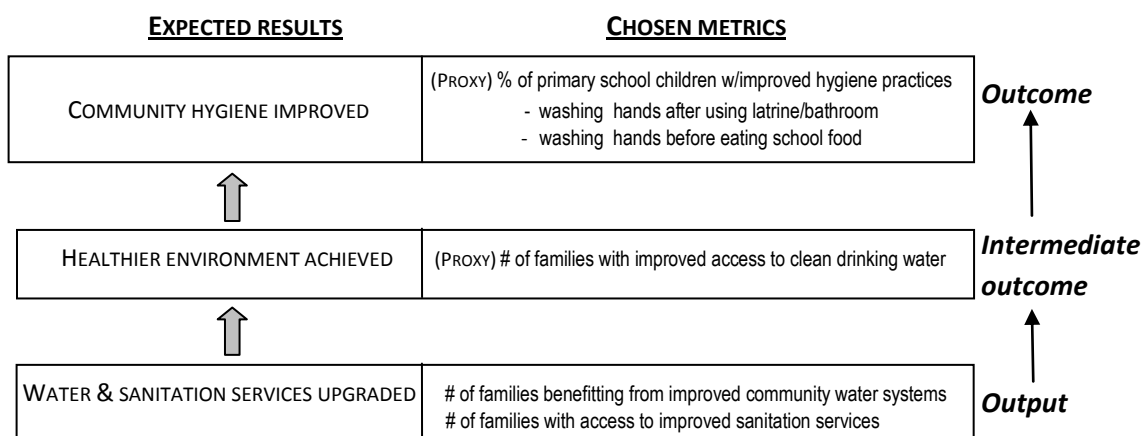


Figure E3: Example of BCDP reconstructed Results Chain in Sanitation

22. There remain five (5) outputs implicit in metrics identified in the *Final List* for the base-line studies that could not be cogently associated with a clearly defined results sequence in the Project. Hence, they are not included in the reconstructed results chains herein presented. Yet, actual measurements of these indicators are separately reported on in this study, because they represent clear contributions to the batey communities. They are: (i) # of families with improved housing; (ii) # of bateys with access to an emergency-safe structure; (iii) # of bateyes with updated emergency plan; (iv) # of batey schools with updated emergency plans; and (v) # of residents receiving support in income generation activities. Now, based on the reconstructed individual results chains, figure E4 summarizes the integrated development logic (*project theory*) postulated by the present researchers for the evaluation of the BCDP. This retrofitted Results Chain, together with the full set of performance metrics and some ex-post identified effectiveness risk factors, is reflected in the Project *LogFrame* matrix of the following page.

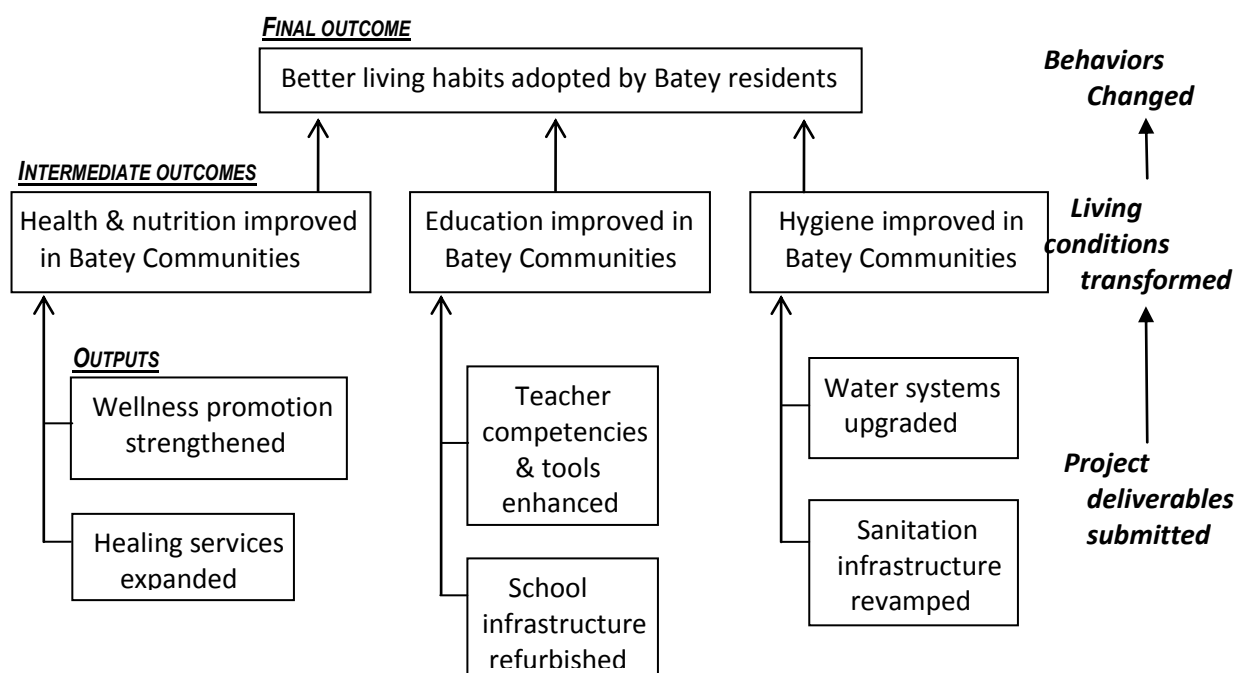


Figure E4: Integrated Results Chain. (Retrofitted development logic: hypothesis for BCDP evaluation)

BCDP LOGICAL FRAMEWORK MATRIX (partial, reconstructed)

Expected Results	Performance Indicators	Means of Verification	Assumptions
GOAL (Expected impacts) Quality of life raised for Batey residents	n/d ²⁵	n/d	<u>From goal to sustainability</u> n/d
PURPOSE (Terminal expected outcome) Better living habits adopted by Batey residents (Expected intermediate outcomes) 1. Health & nutrition improved in Batey Communities 2. Education improved in Batey Communities 3. Hygiene improved in Batey Communities	(All metrics refer to the batey communities relevant to the BCDP) - % of care givers reporting three key child health & nutrition practices - % of infants under 6 months exclusively breastfed in the past 24h. - % of women ages 15-49 reporting two key reproductive health practices - (Proxy) % of adults identifying two STI and HIV/AIDS prevention practices - % of primary school children improving hygiene practices (2 indicators) - % of children under 5 with diarrhea in the last two weeks - % of students enrolled in 4 th grade that pass 4 th grade 1.1 % of women ages 15-49 receiving one PAP smear test in the past year 1.2 # of children under 5 receiving vitamin A from USG programs 2.1 (PROXY) # of children enrolled in pre-school 2.2 (PROXY) % of enrolled primary schools students attending 2.3 # of students (6-14 years old) enrolled on USG supported extra-curricular programs 2.4 # of youth & adults participating in USG supported vocational or informal education 3. (PROXY) # of families with improved access to clean drinking water	- Selection of a group, randomly assigned among the population not to be intervened by the Project (Control group) that is statistically comparable with the population to be intervened (Treatment group). - Independent field measurement of indicators at the baseline and project-end points, on both Treatment and Control groups. - Statistical significance test applied to empirical measurements to discard chance or random errors	<u>From purpose to goal</u> - The target population sustains the effort of organization and self-management in the critical areas of project intervention.
COMPONENTS (Expected outputs) 1. Health 1.1 Healing services expanded 1.2 Wellness promotion strengthened 2. Education 2.1 School infrastructure refurbished 2.2 Teacher competencies & tools enhanced 3. Sanitation 3.1 Sanitation infrastructure revamped 3.2 Water systems upgraded	1.1.1 % of children under 12 months of age who have received DPT3 1.1.2 # of children under 10 years of age who are de-wormed by USG-funded program 1.1.3 # of adults tested for TB 1.1.4 # of people benefitting from USG supported health service 1.1.5 # of bateys with functional MISPAS reference systems 1.2.1 # of community health promoters trained (4 indicators) 1.2.2 # of community health promoters equipped 1.2.3 # of adults tested for TB 2.1 # of classrooms rehabilitated and equipped 2.2.1 # of teachers from participating batey schools trained (5 indicators) 2.2.2 # of classrooms with improved didactic materials 2.2.3 # of administration and education officials trained 3.1 # of families with access to improved sanitation services 3.2 # of families benefitting from improved community water system	Project monitoring system's reports	<u>From components to purpose</u> - The target population has a minimum of predisposition to cooperate in the achievement and submission of Program outputs - Critical cultural and socio-economic resistance factors in the batey community do not excessively impede changes - Social permeability between groups does not generate extraneous factors influencing attitudinal or behavioral changes - Unexpected cross-effects between project components do not impede or retard development outcomes

²⁵ Not defined. Proxy indicators that would be appropriate to measure this type long term transformation (*impact*) are, for instance, the *rate of infant mortality rate*, or the *rate of post-elementary graduation*, in the batey communities. However, no such indicators were identified, or measured, by the BCDP, for inclusion in the evaluation.

Note: 5 additional indicators of outputs outside not included, because they fall outside the conceptual realm of the above Results Chain. They are reported on separately.

ANNEX III: RESEARCH METHODOLOGY

October 2013

This is a Methodology Appendix for The Development Effectiveness Evaluation Study of the USAID - BATEY COMMUNITY DEVELOPMENT PROJECT performed by STC and MUDE.

Method

Design

The design to evaluate this program had two modalities: Quantitative and Qualitative. The Quantitative modality includes the frequency and percentage indicators contemplated in the SOW regarding infrastructure improvements in housing, water and sanitation, as well as education and community health interventions.

Then, the Qualitative modality includes the measures of the impact that the program have had in the community participation and the developing of income generating activities, as well as in the sustainability of the social programming through the partners' organizations.

QUALITATIVE MODALITY

The impact of the program in the community participation as well as the developing of income generating activities was measured through in depth interviews with community leaders. The sustainability of the applied social programming was also measured through in depth interviews with officers from the partners' organizations: MUDE, MISPAS, INFOTEP, etc.

QUANTITATIVE MODALITY

1. FIRST PHASE (2009 – 2013) 8 Communities

To evaluate the original program in the original eight San Pedro de Macorís (SPM) – Hato Mayor bateys we used a Quasi-experimental design with pre and posttest, to compare two groups, Treatment and Control. This design corresponds to the number 4 experimental design from Campbell and Stanley (1963), but in this case, we considered it as a Quasi-experimental design, due to the lack of a real controlled randomization at the formation of the groups. Actually, what were assigned to the Treatment and Control groups were the communities (Bateys), not the subjects.

These authors considered that this design has a limitation in its external validity because of the possible interaction between the testing and the treatment. This limitation does not

apply in this case because the “testing” consisted only in the collection of public and personal data and were not done on the same subjects but on randomly selected subjects in the communities.

The eight bateys identified as the Treatment Group were Don Juan, Cachena, Prudencio, Consuelito, Jalonga, Doña Lila, Experimental and AB-4 (including batey AB-6).

The four bateys identified as the Control Group were Olivares, Alejandro Bass, Chicharrones and Los Platanitos. This Control Group was intended as such for both the First and Second Phases of the project.

GEODATA was commissioned to perform the Baseline and Final studies for the bateys on the Treatment Group and the baseline for the bateys on the Control Group. We were asked to perform the Final study for the Control Group.

Our task evaluating this quantitative modality was to compare the information gathered on the Baseline studies from GEODATA as an external criteria, with the information from the Final studies, in order to validate the reports from STC on the indicators of the program.

2. SECOND PHASE (2011 – 2013) 4 New Communities

In 2010, 4 new bateys were added to the project, also in the SPM area: Euskalduna, Margarita, Paraiso and Victoria.

The same comparative design from the First Phase will be used to evaluate this Second Phase. As mentioned, the same Control Group will be used for both phases in this geographical area.

3. Verón – Bávaro

Two new areas were added to the project, also in 2010, with limited interventions: Verón – Bávaro (VB) and Bahoruco.

The intervened or Treatment communities in VB were Villa Playwood and Hoyo de Friusa. Kosovo, Matamosquito and Macao are considered The Control Group communities.

These five communities were included in a study performed with the support of Punta Cana Foundation, which is an association that includes the area’s tourist hotels. This was considered as the baseline study for both the Treatment and Control Groups.

As contractors, we were commissioned to perform the Final study for the Control Group in this area, but there was not a Final study for the Treatment Group in VB.

For this reason, it was impossible to apply our described number 4 Quasi-experimental analysis design and we had to limit to simple one to one comparisons among the Baseline for the Treatment and Control group, and the Baseline for the Control Group, as well as compared them with information reported by STC.

4. Bahoruco

Starting also in 2010, limited interventions were performed in the Bahoruco zone, including the communities Batey 3 and Batey Mena as the Treatment Group for the zone. As the Control Group were designed the bateys Los Blocks de Mena and Mena Abajo.

We cannot use the same design as described for the other areas because there was not a Baseline study for these bateys. This was confirmed by one of the officers from MUDE, which was working in the Bahoruco zone. In addition, there was not a Final study for the Treatment Group.

In this case, we only have a Final study for the group designated as the Control Group which was performed by us. Here we can only compare the information gathered in this single study with the information reported by STC.

Validation Variables

To validate our design analysis we have had to demonstrate the homogeneity of the samples to be compared. The different communities to compare must be similar for the comparisons to be valid. For this reason, we selected several Validation Variables (VV) to perform these comparisons.

To perform this validation we choose several variables as the sex, the educational level and the occupation of the respondent; the total of members of the household; the kind of lightning, floor, walls, and roof of the house; and the kind of fuel used in the kitchen.

Another important validation for this program evaluation is the comparison of the Control Groups with respect to the program intervention.

It is necessary to demonstrate that the Control Groups were not intervened by the program, not to obscure its effect. If a group designed as a control effectively received an important intervention of the program activities, the comparison will show a weak effect of the program only because of lack of experimental control. If the control group is really not intervened by the program, the comparison will show the real magnitude of the effect of the program.

For this part of the validation, we selected the questions included in the Final studies on knowing of the existence of the project and having benefited from its activities.

Independent Variables

The independent or explicative variables (IV) were the following:

Type of Group: Treatment vs. Control

The bateys in the SPM area that were intervened for the different components of the program were designated as Treatment Group. Other comparable and often very close bateys were designated as the Control Group to be compared by the end of the program and to verify the effect of the treatment or intervention through the Pretest – Posttest period (Baseline – Final Study).

The Baseline studies were done to record the situation of the communities before starting the intervention program. This information was gathered to compare it with the final result of the community after completion of the program.

The general comparison of the Main Effect of the Type of Group thorough the Pretest (Baseline) – Posttest (End Program or Final) interval has the aspect of the following table, with the cells to place the measurements of the Dependent Variables (DV):

Group	Baseline	End Program
Treatment		
Control		

Age

As there were program indicators exclusively directed to children or youth or adults, as hygiene habits, training participation, and testing for TB respectively, we will be making the same basic comparison of our design analysis for each of these age levels separately.

Sex

Although Sex was planned originally as an independent variable, to compare the differential effect of the program on both sexes, the database from the baseline and final studies for neither of the groups permitted to disaggregate the information related to this variable.

Effectively, the questionnaire used in the studies do not ask for the sex of the persons participating in training, the children being health treated, the children with diarrhea, with

severe cough, with vitamin A, dewormed, and enrolled and attending school. So, the database only can offer the frequency of persons and children included in these activities.

The sex of the respondent was registered in most cases; nevertheless, we have found at least one important error regarding to the sex related data. In the baseline for the Control Group in the SPM area, in the section of the questionnaire directed exclusively to women in reproductive age (15-49 years), there is a question on having a PAP test in the last year. In that database we have found 60 men having had a PAP test. Naturally, before calculating our indicator we have proceeded to correct this error.

Dependent Variables

The DV in this evaluation study are the indicators of the efficiency of the program:

1. Health Area: Improved maternal child, reproductive health, STD-HIV/AIDS prevention, TB prevention and treatment and hygiene among batey residents.

- 1.1 # of people benefitting from USG supported health services
- 1.2 % of children under five with diarrhea in the last two weeks
- 1.3 % children under 12 months of age who have received DPT3
- 1.4 % of children under than 5 years of age who received vitamin A from USG supported programs
- 1.5 % of children under 10 years of age who are de-wormed by USG-funded programs
- 1.6 # of Community Health Promoters trained in:
 - Maternal/newborn health
 - Child health and nutrition
 - Reproductive health/family planning
 - TB management *
- 1.7 # of Community Health Promoters equipped
- 1.8* # of bateys with functional MISPAS reference systems
- 1.9 % of care-takers who report at least three key positive child health and nutrition practices
- 1.10 % infants under 6 months of age exclusively breastfed within the past 24 hours
- 1.11 % women between the ages of 15 and 49 receiving at least one PAP smear test in the past 12 months
- 1.12 % of women between the ages of 15 and 49 who can report at least two key reproductive health practices
- 1.13 % of heads of households who identify at least key two STI and HIV/AIDS prevention practices
- 1.14 # of adults tested for TB
- 1.15 % school-children who wash their hands:
 - After using the school latrine/bathroom

- Before eating the school breakfast/snack/lunch

2. Education Area: Increased access to and improved quality of primary school education and extracurricular educational services.

- 2.1 # classrooms rehabilitated and equipped
- 2.2 # of classrooms with improved didactic materials
- 2.3 # of teachers from participating batey schools trained in:
 - Literacy
 - Math
 - Computer use
 - School Health and Nutrition
 - Other pedagogical improvements/practices
- 2.4 # of administrators and education officials trained
- 2.5 % of enrolled primary school students attending
- 2.6 % of students enrolled in 4th grade that pass 4th grade
- 2.7* % of students in single grade primary schools in 1-4 grade who read at grade level
- 2.8 # of children enrolled in pre-school
- 2.9 # of children and youth (6-14 years old) enrolled in USG supported extra-curricular programs
- 2.10 # of youth and adults participating in USG-supported vocational or informal education programs

3. Infrastructure Area: Improved water, sanitation, school, and housing infrastructure and services.

- 3.1 # of families with access to improved sanitation services
- 3.2* # of families benefitting from improved community water systems
- 3.3 # of families with improved access to clean drinking water
- 3.4 # families with improved housing

Enabling Strategy: Participatory Community Mobilization

- 3.5 # of bateys with access to an emergency-safe structure
- 3.6 # of bateys with updated emergency plan
- 3.7* # of batey schools with updated emergency plan
- 3.8* # of residents receiving support in income generation activities

* New indicator added in 2011 for Extension Project.

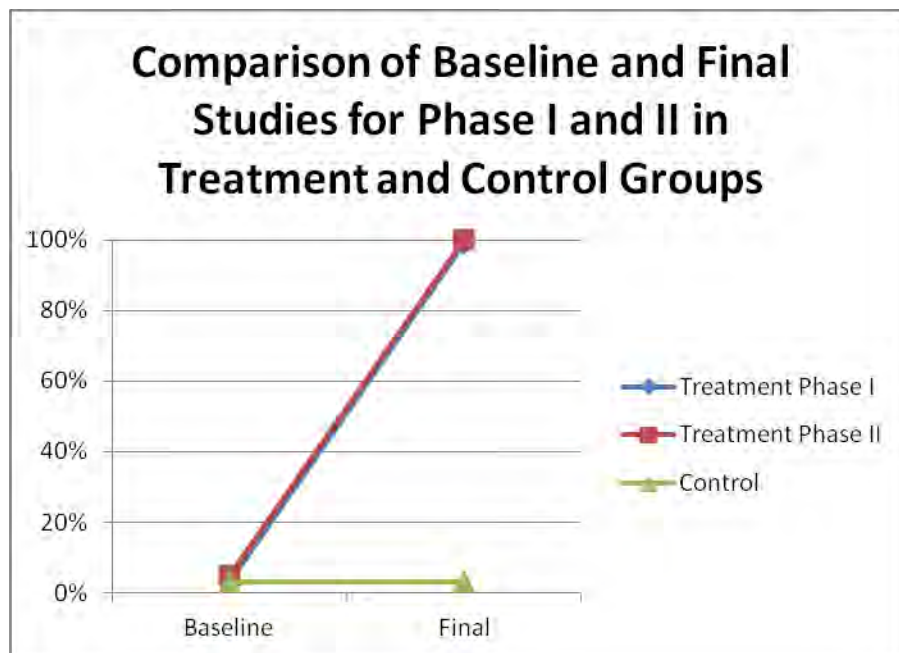
The communities in the SPM-Hato Mayor area were intervened in all these areas and indicators. The communities in the VB area were intervened on 4 health indicators (1.1, 1.3, 1.4 and 1.5) and 1 infrastructure indicator (3.3). The Bahoruco bateys were intervened on the same health indicators as in VB but in 2 infrastructure ones (3.1 and 3.3).

Hypothesis

The expected general hypothesis is that the program has had a significant effect in the Treatment Groups on each of the measured DV, not expecting major improvements in the Control Group.

In the geographical area (SPM) in which it was possible to apply the quasi-experimental design 4 analysis with pretest and posttest, we included the baselines and final studies for both Treatment Groups (Phase I and II), as the baseline and final study for the Control Group.

The ideal expression for this hypothesis is expressed in the following figure:



In the other areas with limited interventions we don't have the Baseline and Final studies for both Treatment and Control groups, but we will be expecting better conditions for the final treatment communities in general.

Samples

As stated in the SOW, and mentioned in the DV section, we only need to collect information from the comparative communities (Control Groups) to complete the set of data needed.

The Samples by community draw by GEODATA for the Baseline and Final studies on the Treatment Groups, and for the Final study on the Control Group, for the Phases I and II (SPM) are shown on the following tables.

TREATMENT BATEYS		
PHASE I	BASELINE	FINAL
Jalonga	53	82
Prudencio	16	35
Cachena	31	39
Doña Lila	16	22
Don Juan	52	94
Consuelito	50	68
AB-4 (+AB-6)	41	56
Experimental	45	65
Total	304	461

FASE II	BASELINE	FINAL
Euskarduna	123	77
Margarita	75	47
Paraiso	78	47
Victoria	66	38
Total	342	209

CONTROL BATEYS	
PHASE I and II	BASELINE
Olivares	29
Alejandro Bass	56
Chicharrones	80
Platanitos	19
Total	184

The samples draw for the baseline study on the Treatment and Control bateys in VB are shown on the next tables.

TREATMENT BATEYS	
VERON-BAVARO	BASELINE
Villa Playwood	252
Hoyo de Friusa	115
Total	367

CONTROL BATEYS	
VERON-BAVARO	BASELINE
Kosovo	139

Matamosquito	64
El Macao	58
Total	261

GEODATA used a proportional procedure to draw probabilistic samples from the population of the bateys, keeping the Confidence Level in 95% and the Confidence Interval below $\pm 4\%$.

As we are using a Quasi-experimental design analysis, to design the sample for the Final study for the Control Group, we do not need to draw a probabilistic sample from a population but a randomized sample for each of the categories of subjects that we will be comparing with respect to the effects of the IVs.

The practical approach to design the numerical aspects of this sample is to perform a Power Analysis (Cohen, 1992).

Power Analysis

With an A Priori Power Analysis we can determine the total minimal sample required to ensure an adequate Size of the Effect and the Power of the analysis.

The Effect Size is a standardized measure of the magnitude of the differences we found. The effect size evaluates the magnitude of the observed differences based on the proportion of variance explained by the factor we are studying.

In other words, two found differences can both be of statistical significance, but one of them can represent a large effect size and the other one can only represent a small effect size.

On the other hand, the power of the analysis is given by the amplitude of the sample used. It comes from subtracting the type II error from 1 ($1 - \beta$). In statistics, the type II error is the probability of accepting the null hypothesis when it is false. It is a false negative. By contrast, a type I error is to reject the null hypothesis when it is true. It is a false positive. This error (α) is the one reported in the .05 significance level.

When comparing two means, to conclude that they are different when in fact they are the same would be a type I error, while concluding that the means are equal when in fact they are different would be a Type II error.

Ideally, a researcher should set a probability of type I error of .05 and a power of .95. We did this Power Analysis to A Priori determine the total minimal sample required to ensure these security standards.

For the calculation of all Power Analyses we used an online calculator from Faul (2013): GPower 3.1.4.

As our DV are measured on a Nominal scale (Categories and Frequencies), the appropriate non parametrical statistic to be used is the Chi-square. For this kind of analysis the Power Analysis takes in account the degree of freedom of the comparison. The greater the value of the degree of freedom, the greater will be the minimal sample size required.

Our basic comparison had to be modified to include the fact that we only have one baseline study for the Control Group of the two Phases in the SPM zone. The table below shows the modified comparison including the total samples used in the different studies.

SPM Zone	BASELINE	FINAL
PHASE I	304	461
PHASE II	342	209
CONTROL	184	200

The Chi-square for this comparison has 2 degrees of freedom. According to the result of the Power analysis, the total sample size needed for an appropriate statistical analysis of our comparisons is 172 cases. As we can see in the above table, the other cells in the SPM comparison totalize 1,500 cases. Then, we are left free to select a random convenience sample of 200 for the Final Control Group, for a total sample size of 1,700 cases, way above the necessary.

The difference of the frequencies in each cell is of no consequence because we will be comparing percentages with the Chi-square analysis. Although the frequencies needed to show the actual effect of the program are shown as indicators in the Annual Performance Report FY 2012 from STC (APR2012), we actually calculate them, when possible, from the GEODATA databases, to validate the APR2012 report.

In the case of VB the degrees of freedom for the comparisons are also 2, and then the total sample size needed is 172. As we can see in the table below for the Control bateys, the baseline study had 261 cases. Again, by selecting a random convenience sample of 100 cases, we exceed the needed 172 ones. We distributed these 100 cases proportionally to the population of the communities, as can be seen in the next table.

CONTROL BATEYS		
VERON-BAVARO	BASELINE	FINAL
Kosovo	139	55
Matamosquito	64	26
El Macao	58	19
Total	261	100

The same Final Control Sample of 100 cases was proposed for the Bahoruco area, also distributed proportionally according to the population of the two communities designed as Control Group, as is shown in the table below.

CONTROL BATEYS	
BAHORUCO	FINAL
Mena Abajo	61
Los Blocks de Mena	39
Total	100

Procedure

The sample of households was randomly selected in the Control Group bateys, excluding the commercial housings. The interviews were directed to key informants (household heads) of 18 years of age or more. The information on reproductive health practices was directed to women from 15 to 49 years of age in each household.

The enumerator personnel were trained in administering a questionnaire adapted from the survey designed and used by STC in the previous studies of the program (See ANNEX 1).

Analysis

After a direct supervision and validation of the data, and the critic and coding of the open questions, the statistical analysis was performed using the SPSS version 20 software. For the calculation of all Chi-square analyses, which implied different databases, we used an online calculator from Preacher (2001).

We found a serious error in the way to do the data entry in the SPSS in the Baseline Control study for the SPM zone. They did a total of 184 interviews for the study but the electronic sheet shows 691 cases. This is because they multiplied the cases by the number of people in the household. In other words, they interviewed one household head in a 4 member household and they registered 4 cases with the same responses of the household head, except on the identification information.

In this situation, if one runs a Frequency procedure in the SPSS, for example on Children under 5 years of age with Diarrhea, one will obtain a result of 128 cases of Diarrhea, but there were only 57 children under 5 years of age in the communities of Control. We had to restructure the whole database to be able to correctly analyze it. After correcting this error we obtained a frequency of Diarrhea of only 25 cases (44%).

All the DV and most of the VV are measured on a Nominal level, e.g.: Categories, which indicates a non-parametric analysis (Chi-square), except for the age of the respondent and

household size, which are measured as a ratio scale (Numerical). In these two cases the planned statistical test is ANOVA, General Lineal Model for 2 IV: Study (Baseline-Final) x Type of Group (Treatment-Control).

ANNEXES 2 to 8 contain all the tables output from the statistical analysis that were used to compose the figures in the text.

Results on Validation

Beneficiary Communities

To demonstrate that the Control Groups were not beneficiaries of the program, we compared the responses of the interviewees in each community to several questions included only in the Final studies.

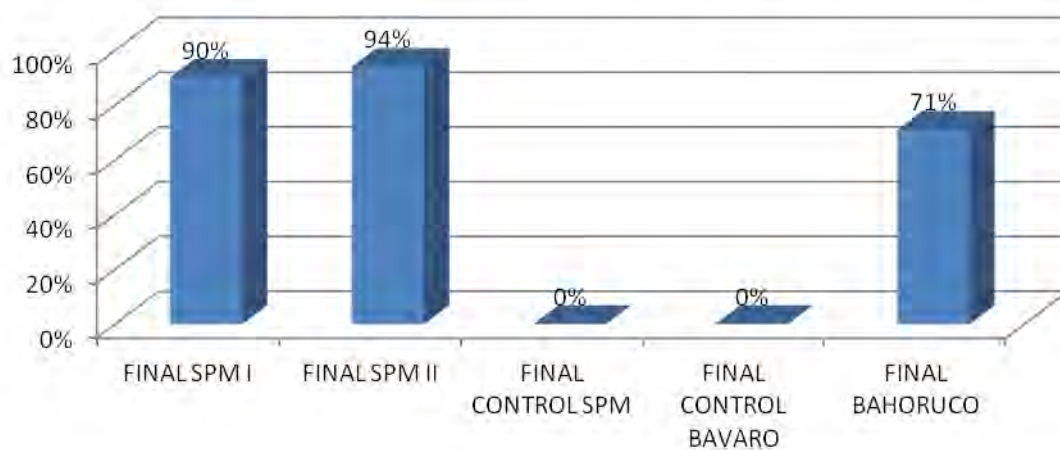
If a particular community reports that they have been beneficiaries of the activities of the program, that particular community must not be considered as a Control Group, to have a valid Treatment-Control comparison. In this way, we can show the real effect of the program on the Treatment communities.

The questions included the general knowledge of the presence and execution of the STC and MUDE project and the general benefits received from it. Also, the particular benefits received from its different activities on Family Planning, STD-VIH/SIDA, Emergency Planning, Health and Environmental Training, Educational Activities, Access to Water and Sanitation Infrastructure, Vaccination times, PAP times, and another activity.

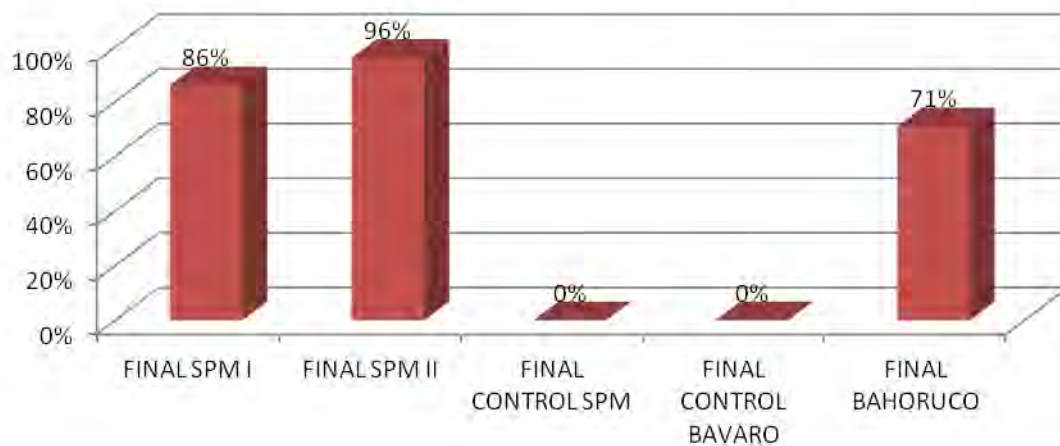
In addition, participants in the SPM zone were asked specifically if they have heard of USAID or AID, and from what country comes from that institution.

The following figures illustrate the differences among the beneficiary communities.

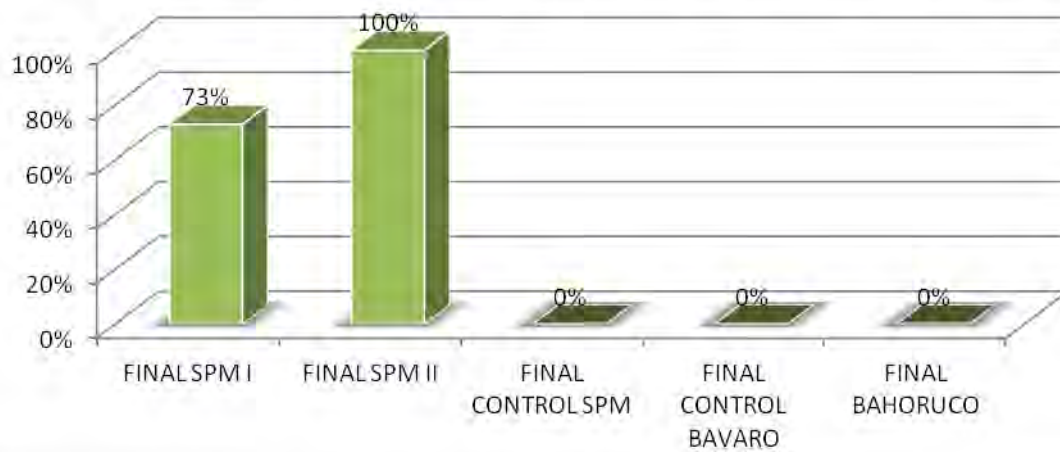
Knows about STC and MUDE project



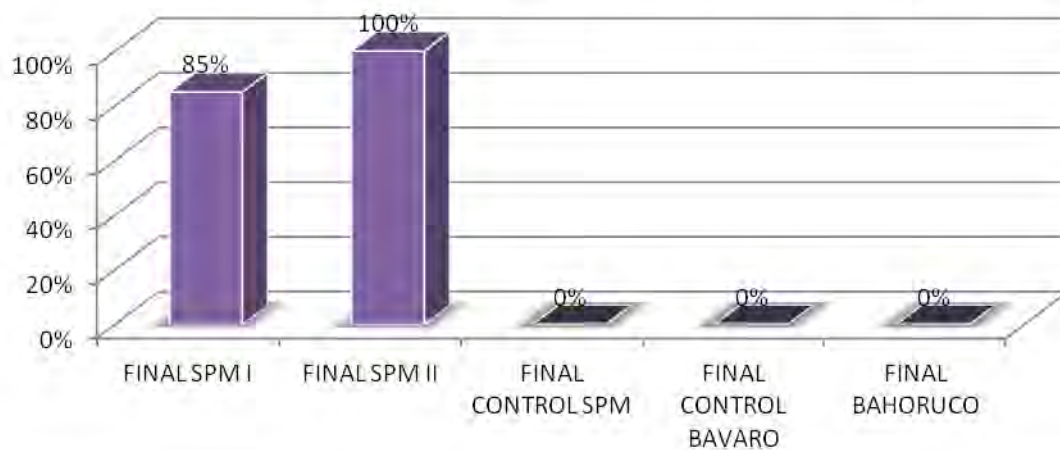
Has benefited from STC and MUDE project



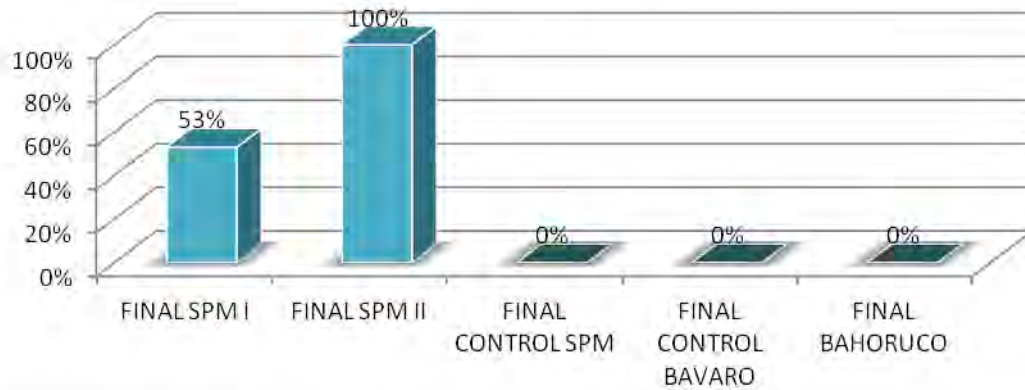
Benefited from Family Planning Final by Zone



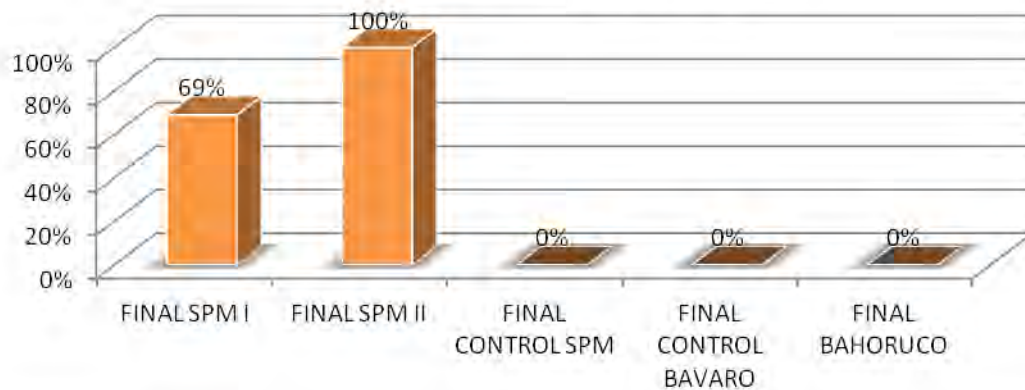
Benefited from STD-HIV/AIDS Final by Zone



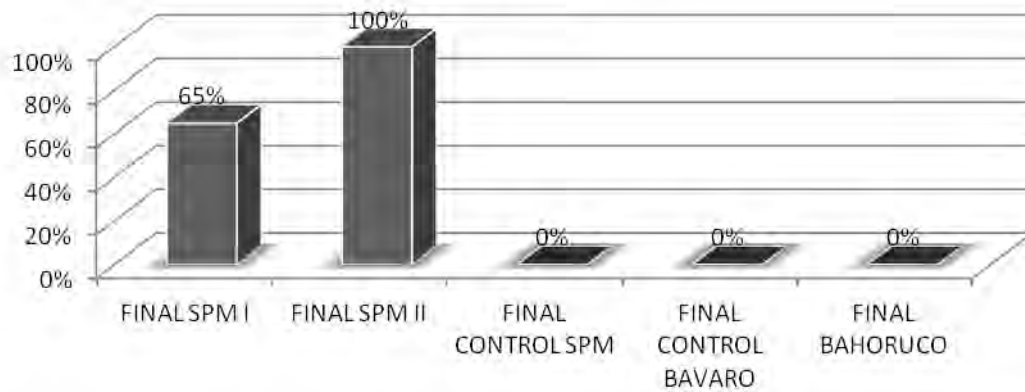
Benefited from Emergency Planning Final by Zone



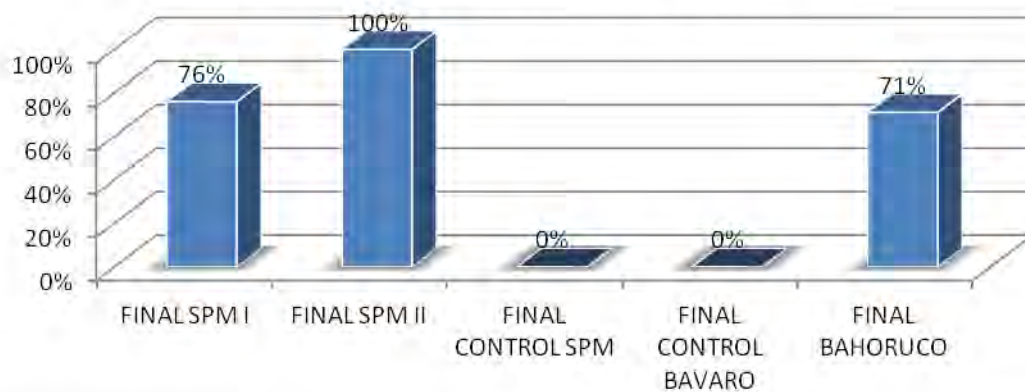
Benefited from Health and Environmental Training Final by Zone



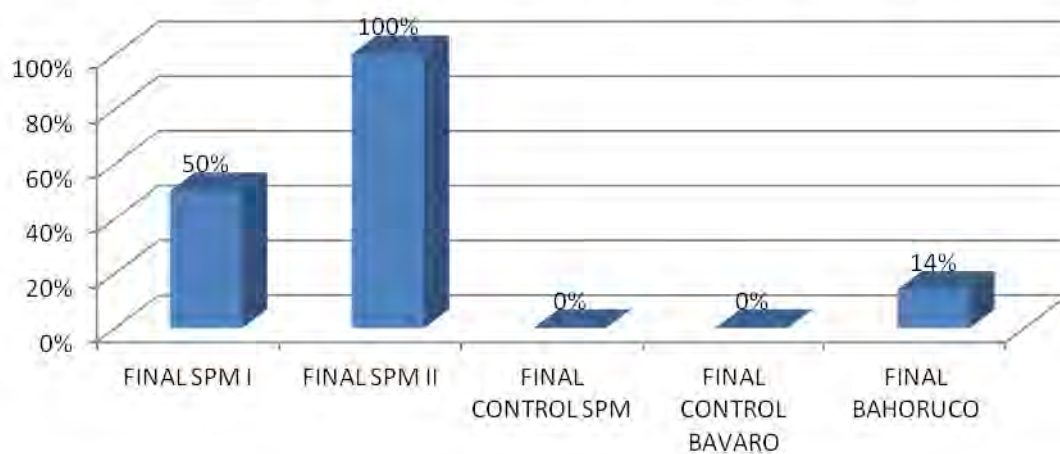
Benefited from Educational Activities Final by Zone



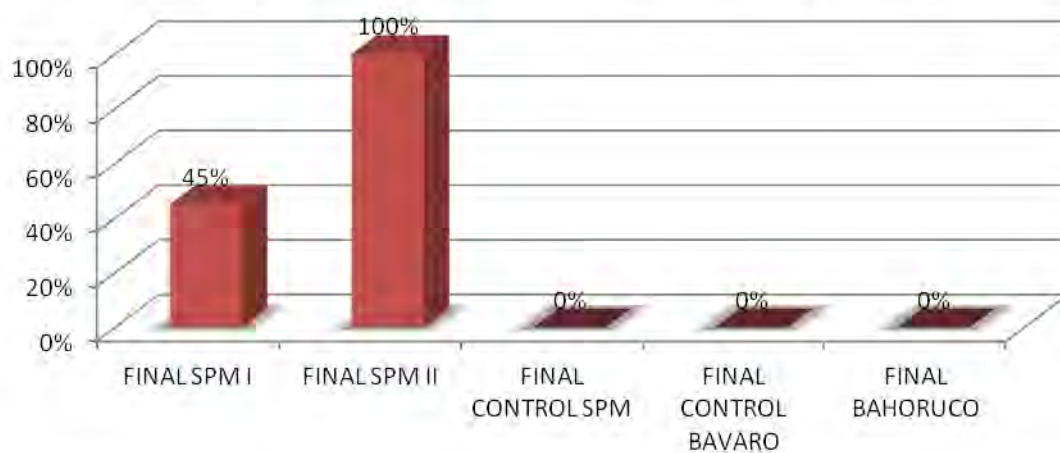
Benefited from Access to Water and Sanitation Infrastructure Final by Zone

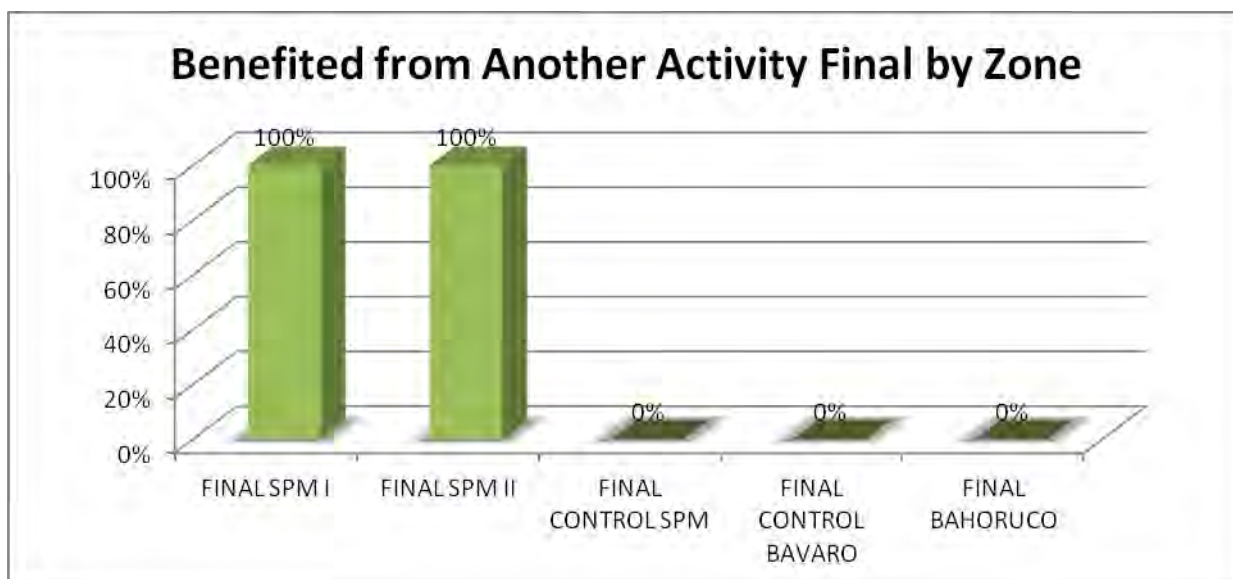


Benefited from Vaccination times Final by Zone



Benefited from PAP times Final by Zone





Consistently, the Final evaluation of the Treatment Groups in the SPM zone shows that there was an effective intervention in those communities, especially in the Phase II of the project. While the percentages of beneficiaries range from 45% (PAP times) to 85% (STD-HIV/AIDS prevention) and to 100% (Another activity) in the Phase I, the beneficiaries reach 100% in every activity in the Phase II.

These results validate the indicator on the reach of the project (1.1). The APR2012 claims that the total population of the Phase I benefited from one or more activities of the project. This is confirmed by the perception of the respondent to the Final studies.

In Phase I 67% of the participants reported having heard of USAID or AID, and 73% of them identified it correctly as an American institution. In Phase II (SPM) 76% of them mentioned having heard of it and 78% of them identified it as an American institution.

The Final Control Groups for SPM y VB were validated as such, because none of them have heard or benefited from the activities of the project.

Nevertheless, 71% of the participants in the Bahoruco study reported that they knew of the presence of the project and that they actually were beneficiaries from, at least, two of its activities: water and sanitation (71%) and Vaccination times (14%).

Discussion on Beneficiary Communities

Officers from MUDE who used to work in the project confirmed that its intervention was not limited to the Treatment Group, but that it reached also the designed Control Group. For these reasons, as the results were very consistent among the interviewees, we ended the data collection in the Mena Abajo and Los Blocks de Mena communities after getting

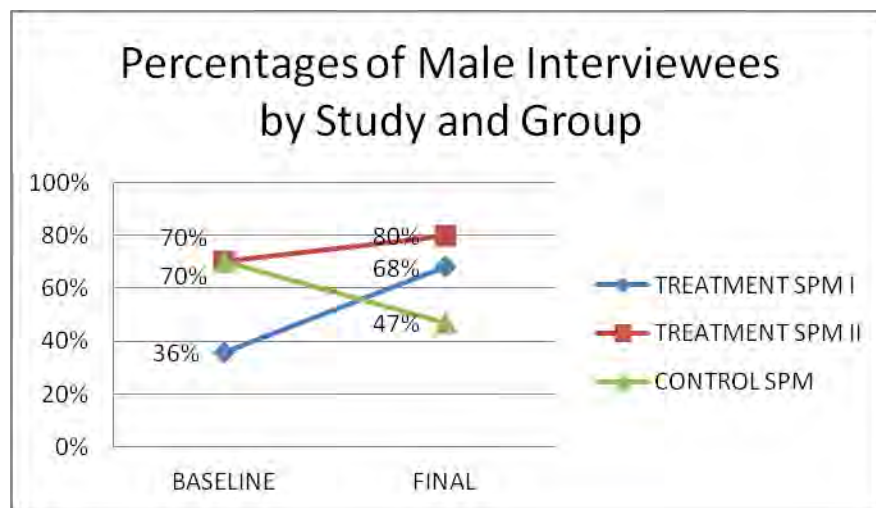
one third of the planned sample, because these communities could not be considered as actual Control Groups.

SPM Validation Variables

To consider our comparisons valid, it is necessary to demonstrate that the communities being compared are similar or comparable, even if they are not exactly the same.

This part of the validation process will deal with the sex, the educational level and the occupation of the respondent; the total of members of the household; the kind of lightning, floor, walls, and roof of the house; and the kind of fuel used in the kitchen.

The results on the sex of the interviewee in each study and group are shown in the table below.



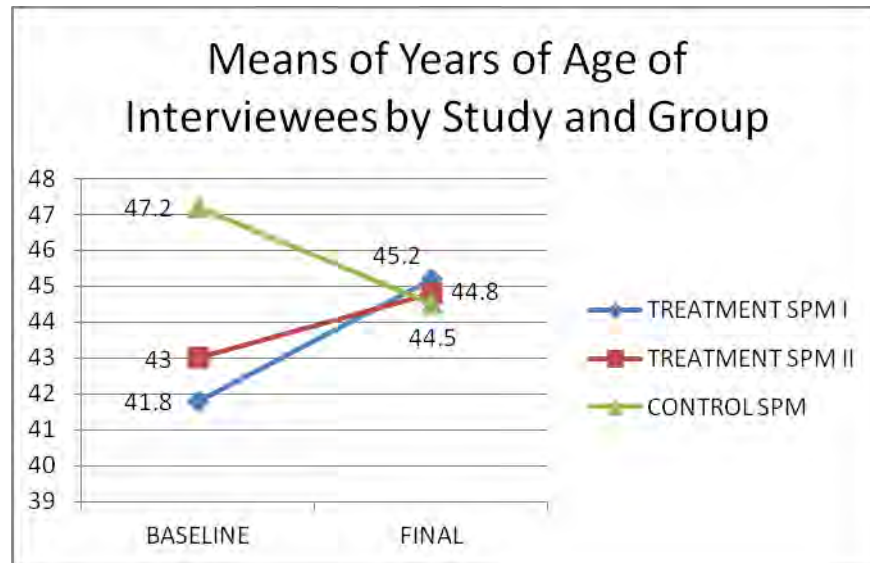
We found some differences among the percentages of males. In SPM I Baseline men are scarce compared to the other groups, except the Final Control, which is close to 50%. Also, the Final SPM I percentage of males (68%) is very much higher than in SPM I Baseline. These differences showed statistical significance (Chi-square (2) = 14.098, $p = .001$, Effect size $w = .25$, medium, Power = .99, Very high).

One factor intervening in these differences could be the interval of time among the different studies. The baseline for Phase I was gathered on 2009, as the baseline for the Control, while the baseline study for Phase II was performed on 2010. The Final study for both Phases was simultaneously done in 2012, while the Final study for the control was performed in 2013.

These can combine with the well-known fact that there is a high rotation of the population in the bateys, producing some of these differences.

Nevertheless, the greatest differences are found among the Baseline studies, and consequently they cannot be influenced by the presence of the program.

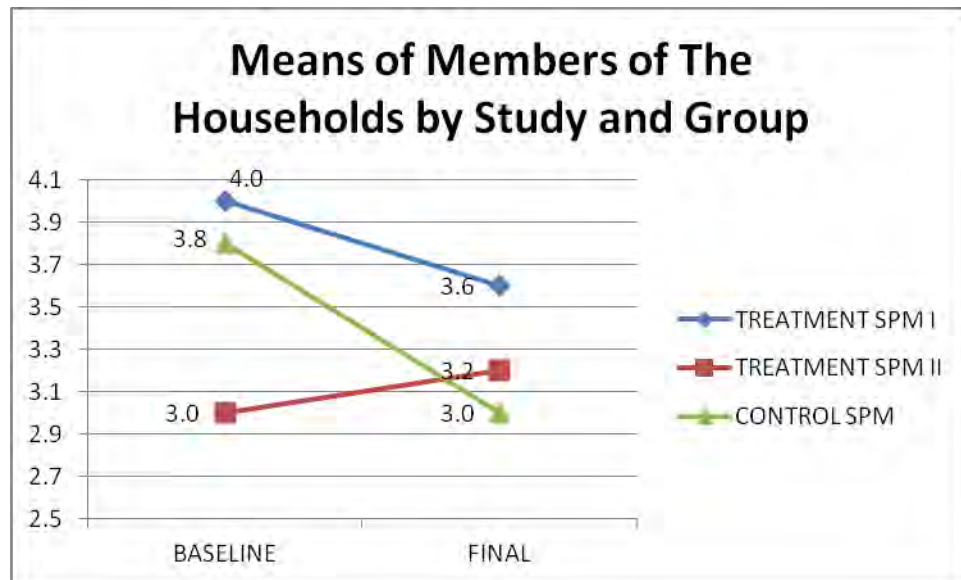
The following figure shows the means of years of age of the interviewees by study and group.



The above figure shows an interaction between the studies and the groups. While every Final study presents almost identical years of age, in the Baseline the Control Group shows an older age than the other groups.

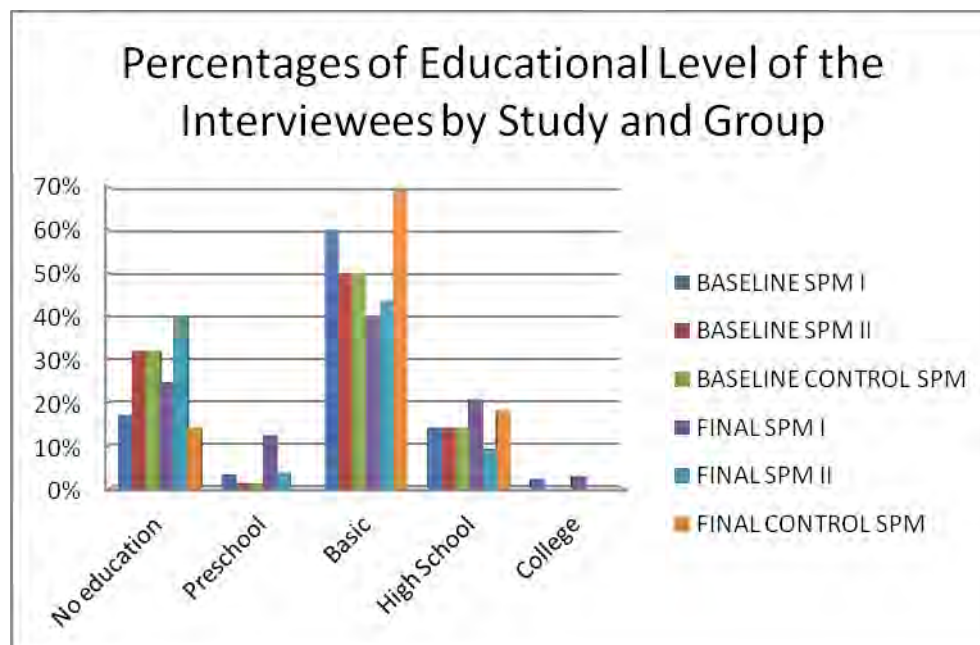
This difference was confirmed as significant with the ANOVA ($F(2/1586) = 4.055$, $p = .014$). Nevertheless, the size of its effect was insignificant ($R^2 = .009$) and the greater difference is 5 years of age, not an important one. The power of the analysis is high (.845) because it involved the entire subjects from all studies.

The next figure contains the means of members of the households by study and group.



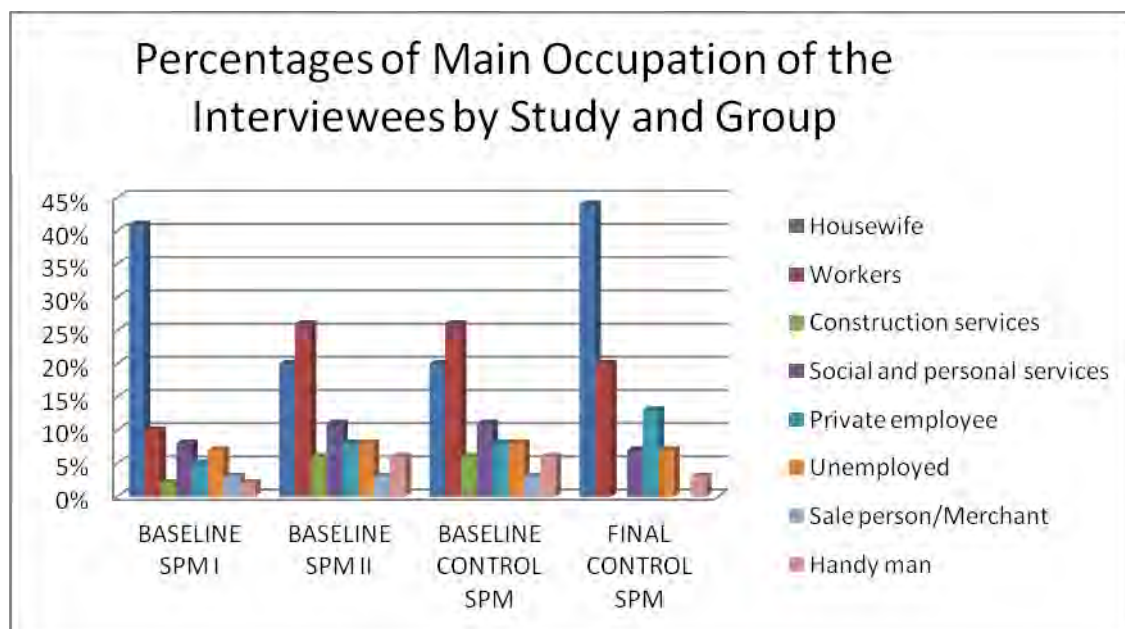
Again, we find some differences due to the interaction between the studies and the groups. Although the highest difference is 1 person per household, in the Baseline, SPM I have a mean of 4 members and the Control 3.8, while SPM II only have 3 members. This small difference was identified as statistically significant by the ANOVA ($F(2/1587) = 5.424, p = .004$) but, once more, the effect size was very small ($R^2 = .031$), explaining only 3% of the variance. For the same reason of a very large sample, the power of the analysis was very high (1).

In the figure below we present the percentages of the educational level of the interviewees by study and group.



This time there are no differences among the studies and groups. In every case the Median educational level of the interviewees is Basic school (52% overall), followed by No education (27%) and High School (15%).

The next figure shows the percentages of the main occupation of the interviewees by study and group.



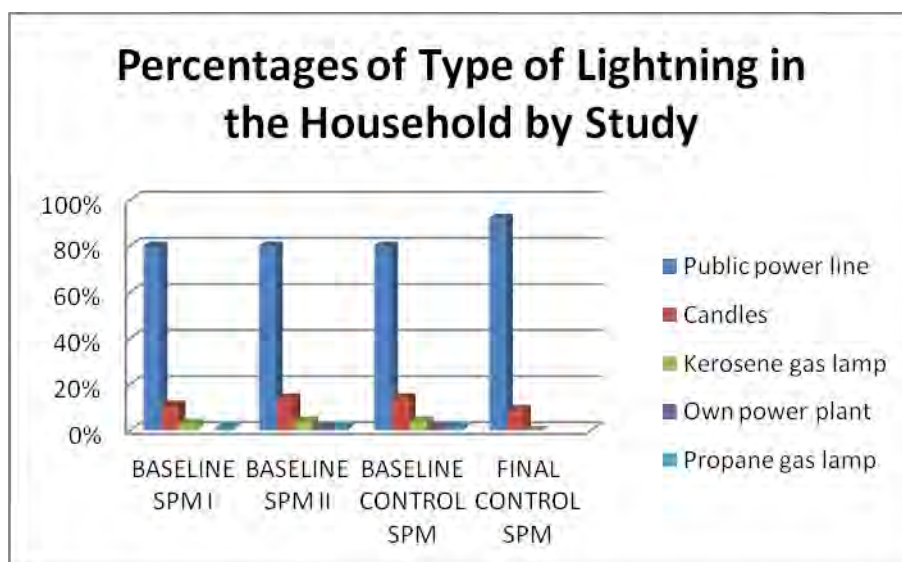
For this variable, it is not possible to make our main comparison by study and group because the occupation of the interviewees was not included in the database from the Final study for Phases I and II in SPM.

Comparing one to one the available data from the different studies we find some differences among the studies for two occupations, Housewife and Workers. Housewives are the majority in the Baseline SPM I (41%) and the Final SPM Control (44%), being Workers the second place in both cases. On the contrary, Workers are in the first place in the Baseline SPM II (26%) and the Baseline SPM Control (26%), being Housewives in the second place for both.

The differences among the studies are significant for the Housewives (Chi-square (3) = 16.344, $p = .001$, Effect size $w = .36$, medium, Power = .94, Very high), and the Workers (Chi-square (3) = 8.341, $p = .039$, Effect size $w = .32$, medium, Power = .67, Medium high).

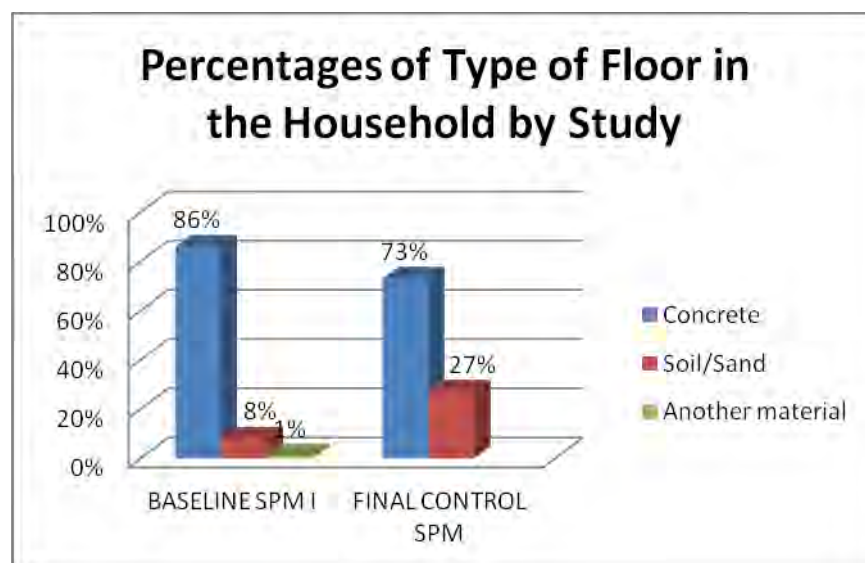
In these two studies the Occupations are more evenly distributed but, anyway, Housewives and Workers are on first or second place in both of them.

The percentages of type of lightning in the household by study can be found in the table below.



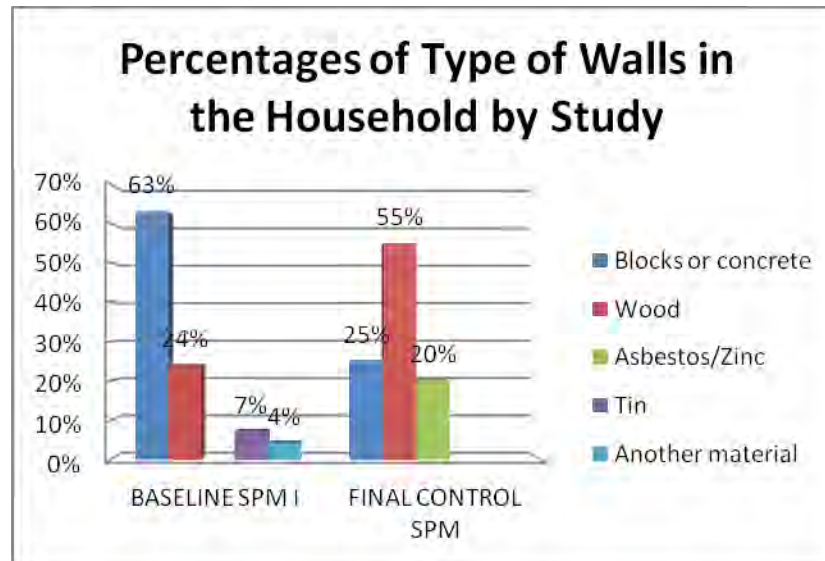
Again, this variable was not included in the database from the Final study for Phases I and II in SPM and we can only make one to one comparisons among the studies.

Nevertheless, this time there are no differences among the studies. The main type of lightning for all communities is the Public power line (83% overall). Very few households use Candles (12%), and the Kerosene gas lamps are scarce (3%).



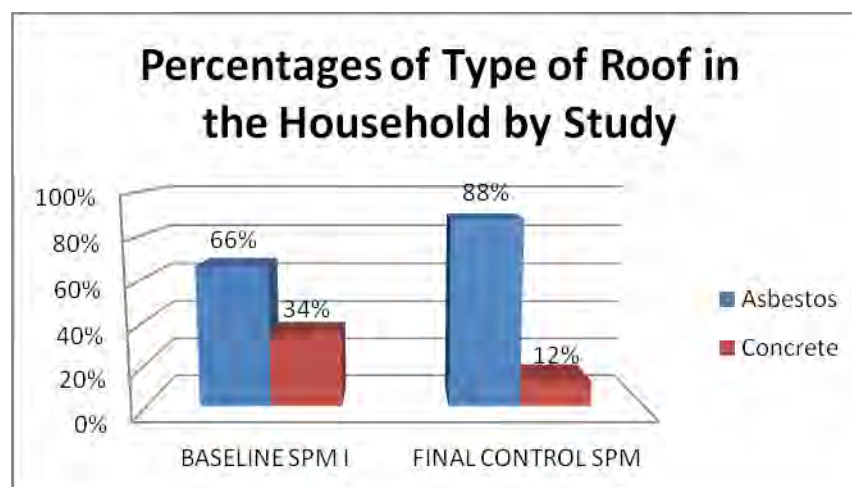
The above table contains the percentages of type of floor in the household by study. This time, these data was recorded only in the Baseline study for the SPM I Treatment Group and the Final study for the SPM Control Group.

As before, there are no differences among the studies. The predominant material for the floor was always Concrete, and only in a few cases it was Soil or sand. The next table shows the percentages of type of walls in the household by study. Again, these data was recorded only in the Baseline study for the SPM I Treatment Group and the Final study for the SPM Control Group.



This time we find an apparent difference between the studies because in the Baseline for SPM I the predominant material of the walls is Blocks or concrete (63%) and in the Final Control study is Wood (55%). The significance of the difference could not be tested because too many of the cells had expected frequencies lesser than 5.

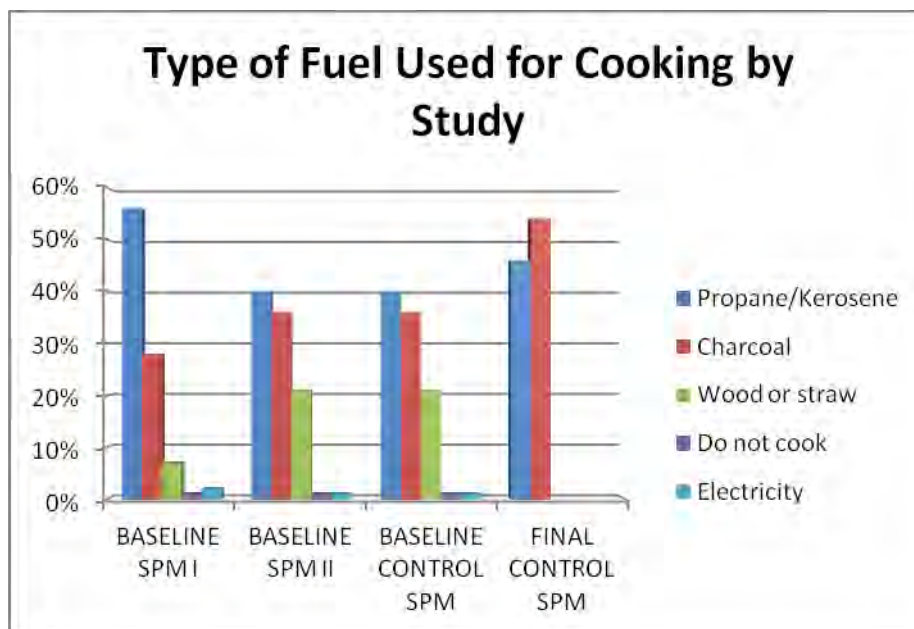
Although this difference could be related to the construction materials available in the communities, these two materials, Concrete and Wood, are in the first or second place in both communities.



The table above contains the percentages of type of roof in the household by study. Once again, these data was recorded only in the Baseline study for the SPM I Treatment Group and the Final study for the SPM Control Group.

There are no differences in the kind of roof in both communities. The first one is always Asbestos/Zinc and the second one Concrete.

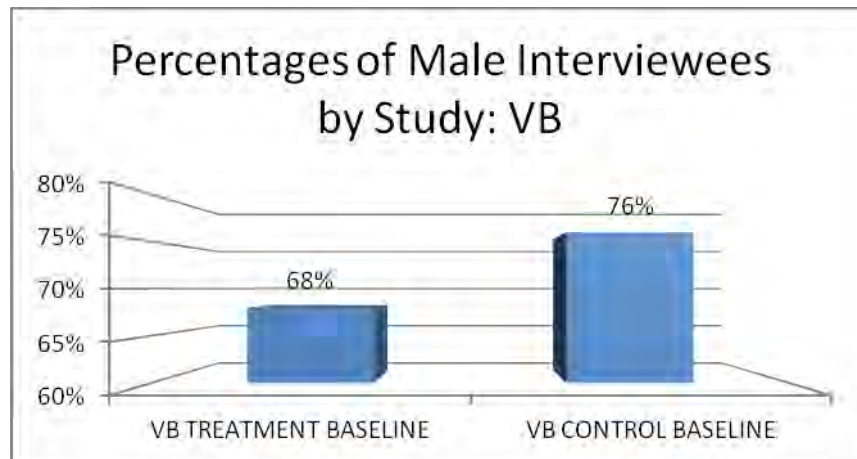
The percentages of type of fuel used for cooking by study can be found in the table below.



The predominant fuel used for cooking is always Propane/Kerosene (46% overall), except for the Final SPM Control Group in which it was Charcoal. Anyway, the Charcoal was in the second place in all other communities, and this difference could be related to the fuel availability in each community.

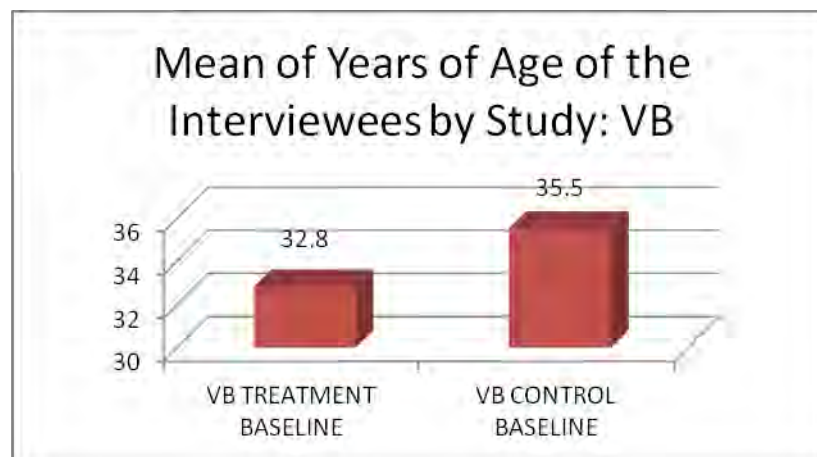
VB-Bahoruco Validation Variables

As we can see in the next figure, in VB we only have data from the Treatment and Control Baseline with respect to the sex of the interviewees.



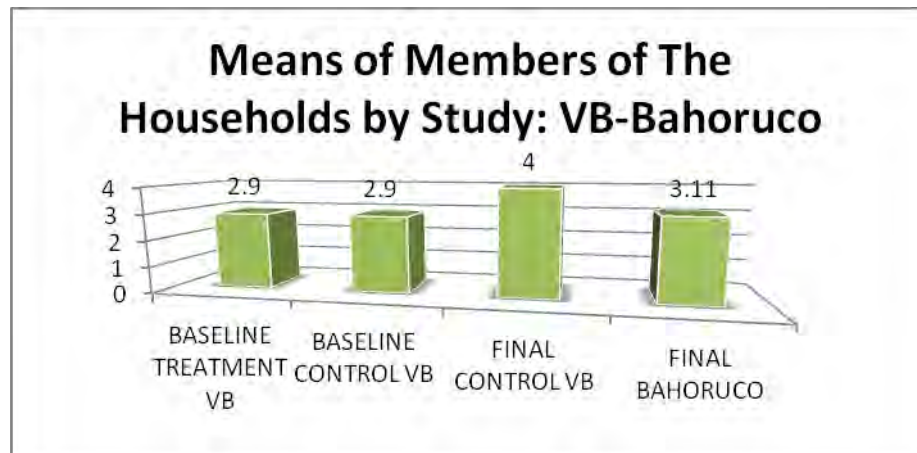
The apparent difference between the studies is not statistically significant (Chi-square (1) = .444, $p = .505$).

The available data on the age of the interviewees from the VB Treatment and Control Baseline is shown on the next figure.



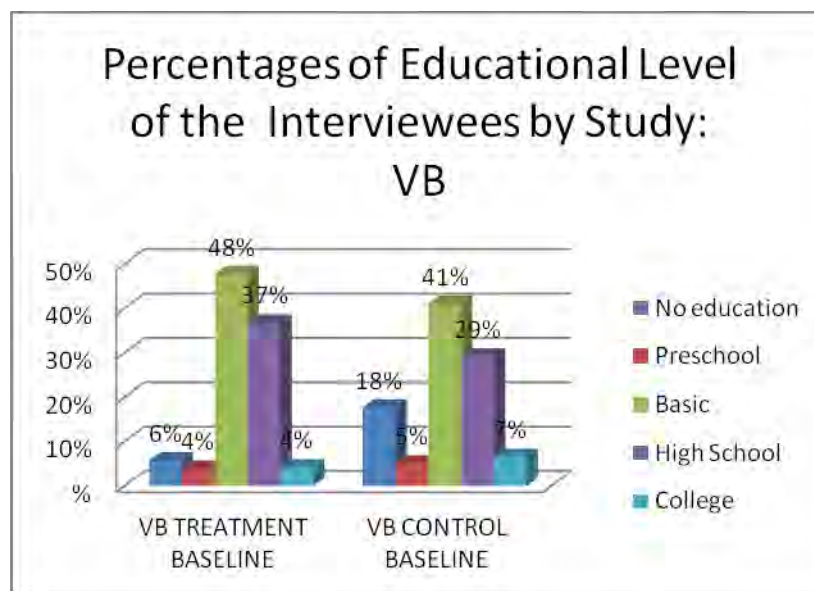
Although the means of age for the interviewees of the two studies is very close, 33 to 36, the difference between them is statistically significant ($t(469) = 2.761$, $p = .006$ (2 tails)). The effect size in this case is medium ($d = .23$) and the power of the analysis is high (.80).

In the case of the total members of the household, we could get data from the studies Baseline Treatment VB, Baseline Control VB, Final Control VB and Final Bahoruco. Their means of members are shown in the figure below.



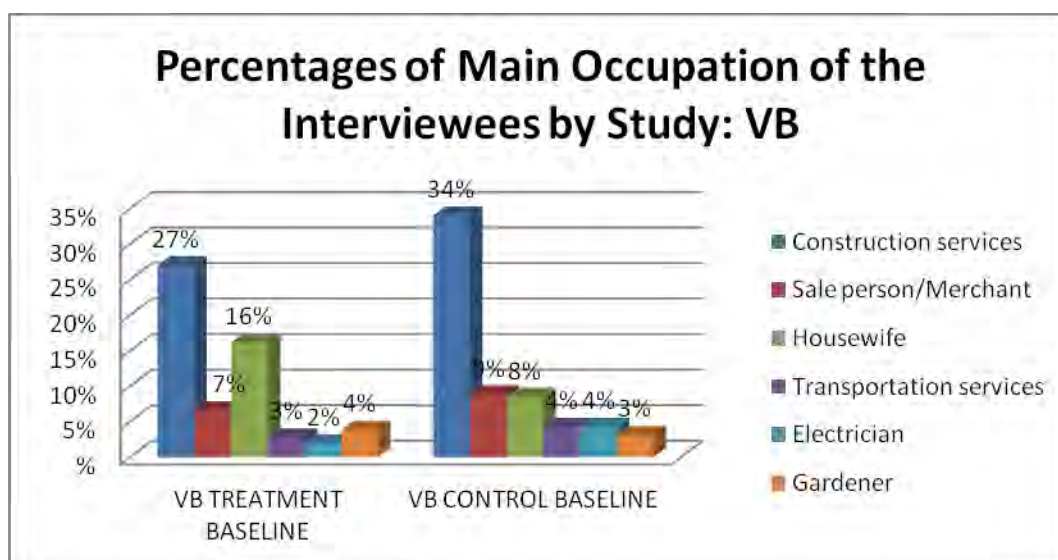
Once more, although the difference is 1 member, the ANOVA found it statistically significant ($F(3/759) = 13.17$, $p = .000$). Nevertheless, the effect size is very small ($R^2 = .049$) because it explains only a 5% of the variance. As is a large sample, the power of the analysis is very high (1).

The percentages of the educational level of the interviewees can be observed in the next figure. Once more, this information was only registered in the Baseline study for the Treatment and Control Group of VB.



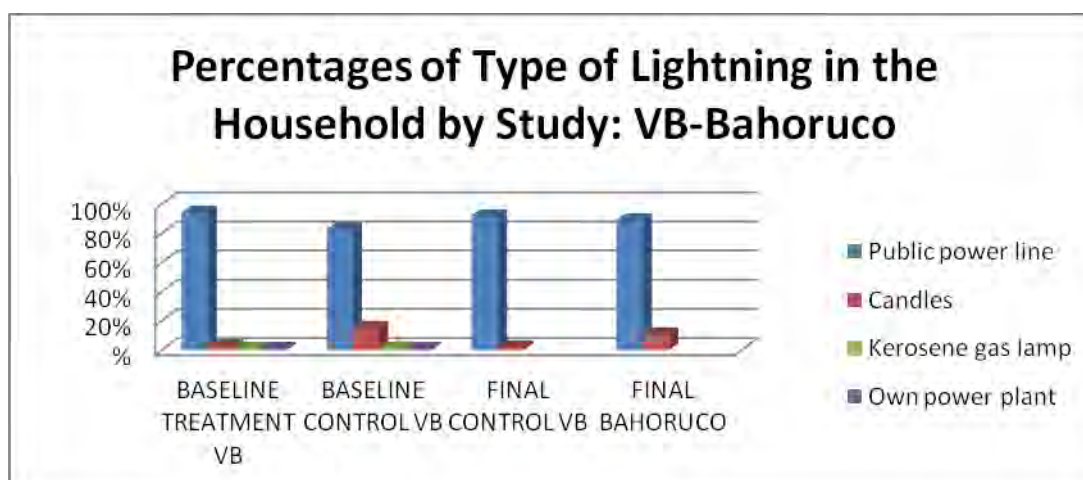
In this case, the median is always in the Basic school education (44% overall), followed closely by High School (33% overall). The No education level is significant in the Control study.

Again, we find information on the main Occupation of the interviewees only in the Baseline study for the Treatment and Control Group of VB. The following figure shows the percentages of the main occupation of the interviewees by study.



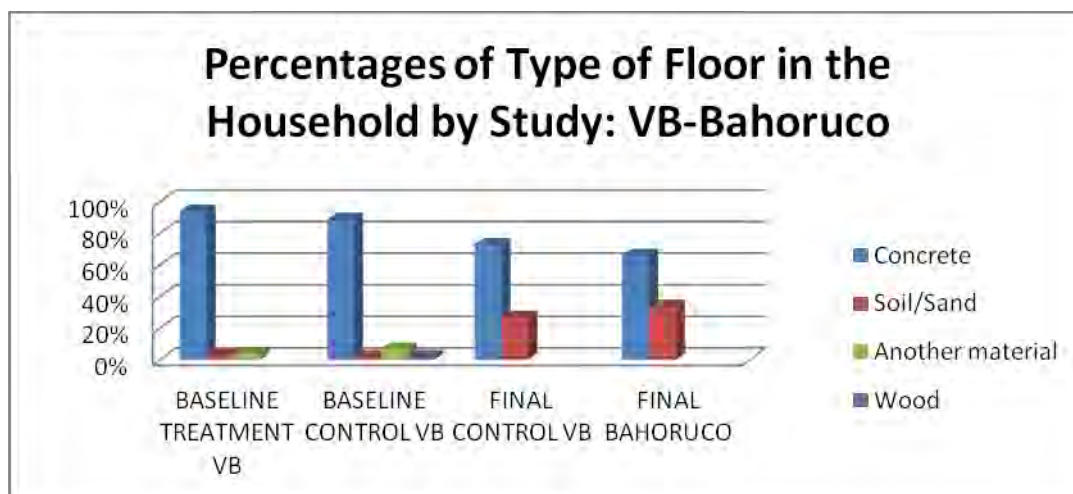
As we can see in the above figure, the main Occupation in both studies in this zone is Construction services (31% overall), followed by Housewife in the Baseline (16%).

For the rest of the validation variables related to the characteristic of the household, we could get data from the studies Baseline Treatment VB, Baseline Control VB, Final Control VB and Final Bahoruco. The figure below contains the percentages of type of lightning in the household by study.



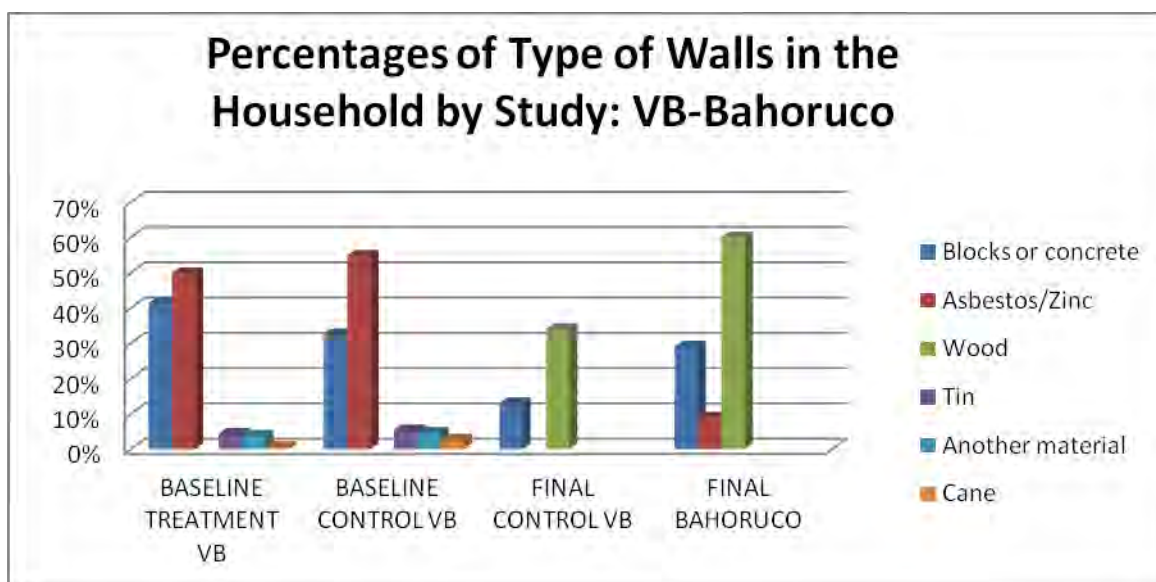
There are no differences among the studies with respect to the type of lightning used in the households. They use almost always the public power line, with a few exceptions that use candles in the Baseline Control VB and the Final Bahoruco.

Percentages of Type of Floor in the Household by Study can be observed in the figure below.



The predominant type of floor in all the studies is Concrete (80% overall), but we find some use of Soil/Sand floor in the Final Control VB (27%) and the Final Bahoruco (34%) studies.

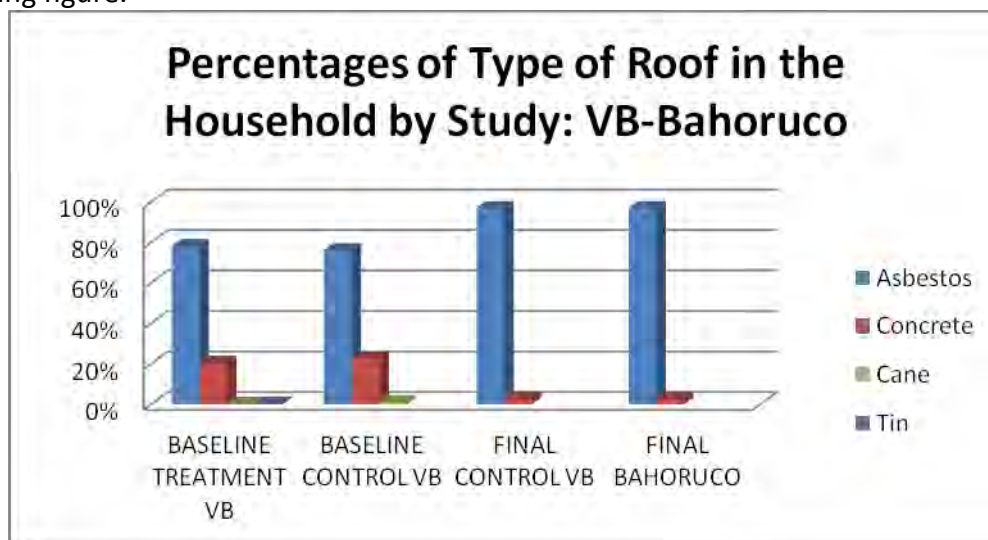
The percentages of the predominant material for the walls of the household by study can be found in the next figure.



Here we found some differences in the materials found by study. In the VB Baseline for the Treatment (50%) and Control Groups (55%), the first place is for Asbestos/Zinc, and the second place is for Blocks or Concrete. Both of these studies were simultaneously performed in 2009.

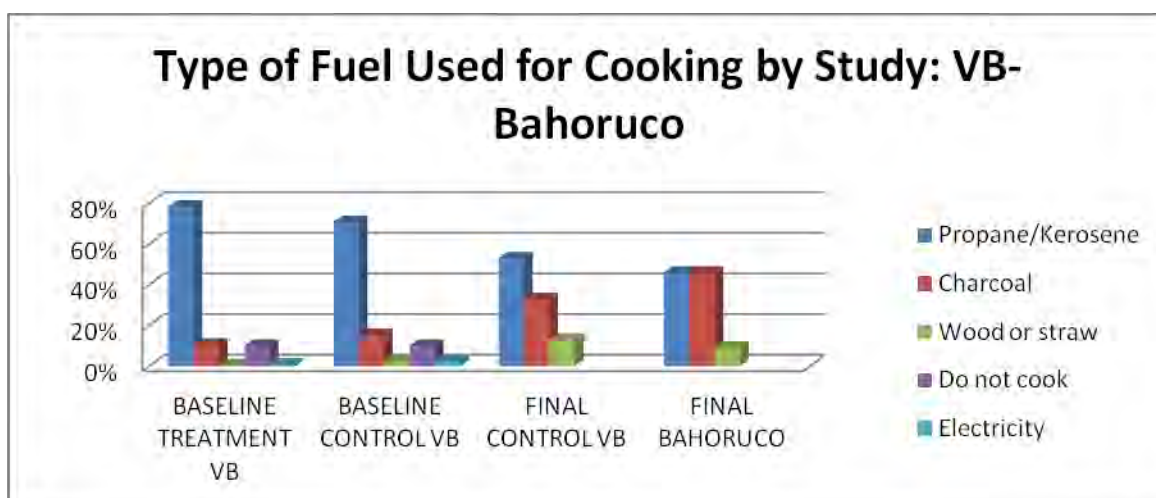
Wood is in the first place for the VB Final Control study (34%) and the Final Bahoruco study (60%). Both studies were made in 2013.

The percentages of type of roof in the household by study can be appreciated in the following figure.



Here we find that Asbestos/Zinc is the predominant material in all studies (87% overall), with a few cases using Concrete in the VB Baseline Treatment and Control studies (22%).

In the next figure we can find the percentages of type of fuel used for cooking in the household by study.



Propane/Kerosene is the main type of fuel used for cooking in almost all studies, with the exception of the Final Bahoruco where it is even with Charcoal (46% each).

Discussion on Validation Variables

Resuming the findings on the VV, we find only one group with a significantly reduced number of males as interviewees (Baseline SPM I). The differences on the ages are statistically significant but small, 42 to 47 in SPM, and 33 to 36 in VB-Bahoruco. The educational level is the same in all communities (Basic school). The occupation of the interviewees changes depending on the communities resources but the predominant are Housewife-Workers-Construction services.

The characteristics of the households are basically the same in all groups, floor of concrete, blocks or wood for the walls, also depending on the community resources, Asbestos/Zinc on the roof, and propane/kerosene or charcoal as fuel for cooking.

In conclusion, the communities are very similar and we find no important differences among them. Consequently, we can consider as valid the comparisons from this evaluation.

References

Campbell, D.T. & Stanley, J.C. (1963). *Experimental and Quasi-experimental Design for Research*. Chicago, Rand McNally.

Cohen, J. (1992). "A power primer". *Psychological Bulletin*. 112 (1): 155–159.
doi:10.1037/0033-2909.112.1.155. PMID 19565683).

Faul, F. (2012) GPower Version 3.1.4,
<http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register>

Preacher, K. J. (2001, April). *Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence* [Computer software]. Available from [%http://quantpsy.org.%">http://quantpsy.org.%](http://quantpsy.org)

Annexes

ANNEX 1: Files CUESTIONARIO CONTROL P 1 y 2
 CUESTIONARIO CONTROL SPM P 3 a 14
 CUESTIONARIO CONTROL VB-BAHORUCO P 3 a 8

ANNEX 2: **File** RESULTS BASE LINE TREATMENT SPM I

Sheets:

SAMPLE

VALIDATION VARIABLES

CHILD HEALTH PRACTICES

PAP

REPRODUCTIVE PRACTICES
 2 HIV PRACTICES
 AGE-EDUCATIONAL
 SCHOOL AGE
 PRE-SCHOOL
 EDUCATIONAL PROGRAMS
 TECHNICAL
 LATRINES
 DOMESTIC WATER
 DRINKING WATER
 HOUSEHOLD MATERIALS
 HOME REPAIRS
 EMERGENCIES

ANNEX 3: File RESULTS BASE LINE CONTROL SPM I & II

Sheets:

SAMPLE
 MEMBERS IN HOUSEHOLD
 VALIDATION VARIABLES
 SEX x RELATIONSHIP
 RELATIONSHIP
 POPULATION AGE
 DIARRHEA
 IMMUNIZATION
 HEALTH PRACTICE
 VITAMIN A
 PAP
 REPRODUCTIVE HEALTH
 2 HIV PRACTICES
 EDUCATIONAL PROGRAMS
 LATRINES
 DOMESTIC WATER
 DRINKING WATER
 HOUSEHOLD MATERIALS
 EMERGENCIES

ANNEX 4: File RESULTS FINAL TREATMENT SPM I & II

Sheets:

SAMPLE
 SEX
 DEMOGRAPHICS
 BREASTFEEDING AGE
 BREASTFEEDING
 DPT3

VITAMIN A
 DEWORMING
 PAP
 REPRODUCTIVE HEALTH
 2 HIV PRACTICES
 LATRINES
 DOMESTIC WATER
 DRINKING WATER
 HOME REPAIRS
 USAID

ANNEX 5: **File** RESULTS FINAL CONTROL SPM I & II

Sheets:

SAMPLE
 MEMBERS IN HOUSEHOLD
 SEX
 AGE
 EDUCATION
 OCCUPATION
 USAID
 HEALTH SERVICES
 2 HIV PRACTICES
 TB TESTS
 SCHOOL HYGIENE
 CHILDREN
 PRACTICES LESS THAN 1 YEAR
 PRACTICES 1 TO 5 YEARS
 DIARRHEA
 DEWORMING
 PRE-SCHOOL
 SCHOOL ATTENDING
 IMMUNIZATION
 VITAMIN A
 BREASTFEEDING
 TECHNICAL EDUCATION
 LATRINES
 DOMESTIC WATER
 DRINKING WATER
 HOUSEHOLD MATERIALS
 HOME REPAIRS
 EMERGENCIES
 REPRODUCTIVE HEALTH
 PAP

ANNEX 6: File RESULTS BASE LINE TREATMENT & CONTROL V-B**Sheets:**

SAMPLE
LATRINES
DRINKING WATER
HOUSEHOLD MATERIALS
AGE

ANNEX 7: File RESULTS FINAL CONTROL V-B**Sheets:**

SAMPLE
MEMBERS IN HOUSEHOLD
USAID
DEWORMING
INMUNIZATION
VITAMIN A
LATRINES
DRINKING WATER
VALIDATION

ANNEX 8: File RESULTS FINAL CONTROL BAHORUCO**Sheets:**

SAMPLE
MEMBERS IN HOUSEHOLD
USAID
INMUNIZATION
VITAMIN A
LATRINES
DRINKING WATER
HOUSEHOLD MATERIALS